

SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: Celine Qian Examiner #: 78710 Date: 1/26/05
 Art Unit: 1636 Phone Number: 2-0777 Serial Number: 09/515363
 Mail Box and Bldg/Room Location: 2A64 Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need. 09/515363

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: Melanoma differentiation associated Gene-5, Promoter & uses
 Inventors (please provide full names): Fisher et al

Earliest Priority Filing Date: 2/29/2000 CRFE

For Sequence Searches Only Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.

Please search for SEQ ID NO: 1 & nucleic acid encoding SEQ ID NO: 2.

Also search SEQ ID NO: 2

seq 1 - 3365NA
 2 - 3371AA
 1035

WMO Vali - CRF
 Date: 1/26/05
 Signature

STAFF USE ONLY

Type of Search		Vendors and cost where applicable
Searcher: _____	NA Sequence (#) <u>1</u> STN _____	
Searcher Phone #: _____	AA Sequence (#) <u>1 + 1</u> Dialog <u>Reverse to NA</u>	
Searcher Location: _____	Structure (#) _____ Questel/Orbit _____	
Date Searcher Picked Up <u>1/26/05</u>	Bibliographic _____ Dr. Link _____	
Date Completed <u>1/28/05</u>	Litigation _____ Lexis/Nexis _____	
Searcher Prep _____ Review Time _____	Fulltext _____ Sequence Systems <u>06p/02p</u>	
Clerical Prep _____	Patent Family _____ WWW/Internet _____	
Online Time _____	Other _____ Other (specify) _____	

This Page Blank (uspto)

Db 181 TATTCACAGACGAGATTTCGGCTATCTCATCTCGTCTTCAGGGCCAGGGTGAAATG 240
Qy 241 TACATCCAGGTGAGCCTGTGTGACATACTGACCTTTCTGCCTCAGAGGTGAAGAG 300
Db 241 TACATCCAGGTGAGCCTGTGTGACATACTGACCTTTCTGCCTCAGAGGTGAAGAG 300
Qy 301 CAGATTCAAGAGACAGTCGCCACCTTCGGGAAACATGCAGGCAGTGAACTGTCTGAGC 360
Db 301 CAGATTCAAGAGACAGTCGCCACCTTCGGGAAACATGCAGGCAGTGAACTGTCTGAGC 360
Qy 361 ACCTTGAGAAAGGAGTCTGGCACCTTGTTGACTCGGGAATTCTGAGAGCCCTCCG 420
Db 361 ACCTTGAGAAAGGAGTCTGGCACCTTGTTGACTCGGGAATTCTGAGAGCCCTCCG 420
Qy 421 AGAACCGGAGCCCTCTGCGCCGCTACATGAACCTGAGCTCACGACTTGCCCTCT 480
Db 421 AGAACCGGAGCCCTCTGCGCCGCTACATGAACCTGAGCTCACGACTTGCCCTCT 480
Qy 481 CCATCGTTGAGAACGCTCATGTATATATCTCCACTGCTGAACCTCCTTCAGCCCACT 540
Db 481 CCATCGTTGAGAACGCTCATGTATATATCTCCACTGCTGAACCTCCTTCAGCCCACT 540
Qy 541 CTGCTGACAAAGCTTCTAGTTAGACGCTCTTGATTAAGTGCATGGAGAGGAACTGTTG 600
Db 541 CTGCTGACAAAGCTTCTAGTTAGACGCTCTTGATTAAGTGCATGGAGAGGAACTGTTG 600
Qy 601 ACAATTGAAGACAGAAACCGGATTGCTGTCAGAAAACAATGAAATGAATCAGGTGA 660
Db 601 ACAATTGAAGACAGAAACCGGATTGCTGTCAGAAAACAATGAAATGAATCAGGTGA 660
Qy 661 AGAGAGCTAATAAAAAAGATTGTGCAAAAAAGAACTGTTCTCTGCATTTCTGAATGTT 720
Db 661 AGAGAGCTAATAAAAAAGATTGTGCAAAAAAGAACTGTTCTCTGCATTTCTGAATGTT 720
Qy 721 CTTGCTCAACAGGAAACAATGAATCTGTCCAAAGATTAAAGGCTCTGATTGCTCAGAA 780
Db 721 CTTGCTCAACAGGAAACAATGAATCTGTCCAAAGATTAAAGGCTCTGATTGCTCAGAA 780
Qy 781 AGCAATGCAGAGATTGAGATTTATCACAAGTTGATGTCCTCAAGTGAAGACCACTT 840
Db 781 AGCAATGCAGAGATTGAGATTTATCACAAGTTGATGTCCTCAAGTGAAGACCACTT 840
Qy 841 CTTTCAACCAAGTTCAAGTCCAAATCTGGAAGAGGCTGCGGCATGAGAAATACTCA 900
Db 841 CTTTCAACCAAGTTCAAGTCCAAATCTGGAAGAGGCTGCGGCATGAGAAATACTCA 900
Qy 901 TCAGATCATCTTTTGCAGATCTTCTGTAGTTTCAAAATCAGACACAAGTTGGCAGAA 960
Db 901 TCAGATCATCTTTTGCAGATCTTCTGTAGTTTCAAAATCAGACACAAGTTGGCAGAA 960
Qy 961 GGAAGTGCAGCTGCTTAGATGAAAGTCTTGACATTAACAGCAATGGGAGTGAATTCA 1020
Db 961 GGAAGTGCAGCTGCTTAGATGAAAGTCTTGACATTAACAGCAATGGGAGTGAATTCA 1020
Qy 1021 GGCACCATGGGAAGTGAATTCAGATGAAGAGATGTGGCAAGAGCATCCCCGAGCCA 1080
Db 1021 GGCACCATGGGAAGTGAATTCAGATGAAGAGATGTGGCAAGAGCATCCCCGAGCCA 1080
Qy 1081 GAACCTCAGCTCAGGCCCTTACCATAAGGAGTTGCCACGCTTGGAAGGGAAGAAT 1140
Db 1081 GAACCTCAGCTCAGGCCCTTACCATAAGGAGTTGCCACGCTTGGAAGGGAAGAAT 1140
Qy 1141 ATCATCATCTGCTCCCTACAGGAGTGAAAAAACAGAGTGGCTGTTTACATTGCCAAG 1200
Db 1141 ATCATCATCTGCTCCCTACAGGAGTGAAAAAACAGAGTGGCTGTTTACATTGCCAAG 1200
Qy 1201 GATCATTAGACAAGAAAGAAAGCATCTGAGCCTGGAAGAAAGTTATAGTTCTGTCAAT 1260
Db 1201 GATCATTAGACAAGAAAGAAAGCATCTGAGCCTGGAAGAAAGTTATAGTTCTGTCAAT 1260
Qy 1261 AAGGTACTGTAGTTGAACAGCTCTTCCGCAAGAGATTCCAACATTTTGAAGAAATGG 1320
Db 1261 AAGGTACTGTAGTTGAACAGCTCTTCCGCAAGAGATTCCAACATTTTGAAGAAATGG 1320

Qy 1321 TATCGTGTATTGGATTAAAGTGTGATACCCAACTGAAAAATATCATTTCCAGAAAGTGTG 1380
Db 1321 TATCGTGTATTGGATTAAAGTGTGATACCCAACTGAAAAATATCATTTCCAGAAAGTGTG 1380
Qy 1381 AAGTCTGTGATATTATATACAGTACAGCTCAAAATCCTTGAAAACTCCCTCTTAACTTG 1440
Db 1381 AAGTCTGTGATATTATATACAGTACAGCTCAAAATCCTTGAAAACTCCCTCTTAACTTG 1440
Qy 1441 GAAATGGAAGAAAGATGCTGGTGTCAATTGTCAAGCTTTTCCCTCATTTATCATTTGATA 1500
Db 1441 GAAATGGAAGAAAGATGCTGGTGTCAATTGTCAAGCTTTTCCCTCATTTATCATTTGATA 1500
Qy 1501 TGTATTCACACCAACAAAGAGCAGTGTATTAATTAATCATCATGAGCATTAATTGATGAG 1560
Db 1501 TGTATTCACACCAACAAAGAGCAGTGTATTAATTAATCATCATGAGCATTAATTGATGAG 1560
Qy 1561 AAGTTGAAAAACAATAGACTCAAGAAAGAAACCAAGTATTCCTTCCCTCAGATA 1620
Db 1561 AAGTTGAAAAACAATAGACTCAAGAAAGAAACCAAGTATTCCTTCCCTCAGATA 1620
Qy 1621 CTGGGACTAACAGCTTCACTGCTGTTGGAGGGGCCAAGCAAGCCAAAGCTGAAGAA 1680
Db 1621 CTGGGACTAACAGCTTCACTGCTGTTGGAGGGGCCAAGCAAGCCAAAGCTGAAGAA 1680
Qy 1681 CACATTTTAAACTATGTGCCAATCTTGATGCATTTACTATTAAACTGTTTAAAGAAAC 1740
Db 1681 CACATTTTAAACTATGTGCCAATCTTGATGCATTTACTATTAAACTGTTTAAAGAAAC 1740
Qy 1741 CTTGATCAACTGAAAAACCAATACAGAGCCATGCAAGAGTTGCCATTCAGATGCA 1800
Db 1741 CTTGATCAACTGAAAAACCAATACAGAGCCATGCAAGAGTTGCCATTCAGATGCA 1800
Qy 1801 ACCAGAGAAGATCCATTTAAAGAGAACTTCTAGAAATTAATGACAAAGATTCAAACTTAT 1860
Db 1801 ACCAGAGAAGATCCATTTAAAGAGAACTTCTAGAAATTAATGACAAAGATTCAAACTTAT 1860
Qy 1861 TGTCAATGAGTCCAAATGTCAGATTTTGGAACTCAACCTATGAACAAATGGGCCATTCAA 1920
Db 1861 TGTCAATGAGTCCAAATGTCAGATTTTGGAACTCAACCTATGAACAAATGGGCCATTCAA 1920
Qy 1921 ATGAAAAAAGCTGCAAAAAAAGGAATCGCAAAAGAACTGTTGTGCAAGAACTTTG 1980
Db 1921 ATGAAAAAAGCTGCAAAAAAAGGAATCGCAAAAGAACTGTTGTGCAAGAACTTTG 1980
Qy 1981 AGGAATGTAATGAGGCCCCCAAAATTAATGACAAATTCGAATGATAGATGCGTATACT 2040
Db 1981 AGGAATGTAATGAGGCCCCCAAAATTAATGACAAATTCGAATGATAGATGCGTATACT 2040
Qy 2041 CATCTTGAACCTTTCTATTAATGAAGAGAAAGATTAAGAGTTTGACATGATGAAGATGAT 2100
Db 2041 CATCTTGAACCTTTCTATTAATGAAGAGAAAGATTAAGAGTTTGACATGATGAAGATGAT 2100
Qy 2101 AGTGATGAGGGTGTGATGATGATGATTTGATGATGATGAAGATGAGATGATTTAAAG 2160
Db 2101 AGTGATGAGGGTGTGATGATGATGATTTGATGATGATGAAGATGAGATGATTTAAAG 2160
Qy 2161 AAACCTTTGAAACTGATGAACAGATAGATTTCTCATGACTTTATTTTGAACAAT 2220
Db 2161 AAACCTTTGAAACTGATGAACAGATAGATTTCTCATGACTTTATTTTGAACAAT 2220
Qy 2221 AAAATGTTGAAAAAGCTGCTGAAAAACCCAGAAATAGAAAAATGAAAAAGCTGACCAATTA 2280
Db 2221 AAAATGTTGAAAAAGCTGCTGAAAAACCCAGAAATAGAAAAATGAAAAAGCTGACCAATTA 2280
Qy 2281 AGAAATACCATTAATGAGCAATATACTAGAGCTGAGGAATCAGCAGAGGAATTAATCTTT 2340
Db 2281 AGAAATACCATTAATGAGCAATATACTAGAGCTGAGGAATCAGCAGAGGAATTAATCTTT 2340
Qy 2341 ACAAAAAACAGACAGAGTGCATATGCGCTTCCAGTGATTAAGTGAAGAAATTTT 2400
Db 2341 ACAAAAAACAGACAGAGTGCATATGCGCTTCCAGTGATTAAGTGAAGAAATTTT 2400

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 27, 2005, 10:33:21 ; Search time 13648 Seconds
(without alignments)
11659.572 Million cell updates/sec

Title: US-09-515-363C-1
Perfect score: 3365
Sequence: 1 gcgcgcgcgcctgagagccc.....aaatgattgttactctg 3365

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3365	100.0	3380	6	CQ776470	CQ776470 Sequence
2	3365	100.0	3380	9	AF095844	AF095844 Homo sapi
3	3338.8	99.2	3372	6	AX300832	AX300832 Sequence
4	3338.8	99.2	3373	9	AY017378	AY017378 Homo sapi
5	2476	73.6	2540	9	BC078180	BC078180 Homo sapi
6	2202.4	65.5	3771	6	CQ777379	CQ777379 Sequence
7	2202.4	65.5	3771	10	AF374384	AF374384 Mus muscu
8	2135.4	63.5	3078	10	AY075132	AY075132 Mus muscu
9	1766	52.5	1776	6	AX714144	AX714144 Sequence
10	1766	52.5	1776	9	AK056293	AK056293 Homo sapi
11	1578.8	46.9	2028	6	CQ729104	CQ729104 Sequence
12	1435.6	42.7	1443	6	AX300841	AX300841 Sequence
13	1402	41.7	2468	10	BC025508	BC025508 Mus muscu
14	1278.8	38.0	1392	6	AR379503	AR379503 Sequence
15	1254.6	37.3	1284	6	AX300838	AX300838 Sequence
16	1193.4	35.5	1204	9	BC007966	BC007966 Homo sapi
17	1186	35.2	2050	10	BC004031	BC004031 Mus muscu
18	803	23.9	1258	6	AX098232	AX098232 Sequence
19	803	23.9	1270	6	AX098236	AX098236 Sequence

20	797	23.7	1630	9	BC046208	BC046208 Homo sapi
21	663	19.7	3692	6	AX098234	AX098234 Sequence
22	663	19.7	3704	6	AX098238	AX098238 Sequence
23	623.2	18.5	95417	9	AC010876	AC010876 Homo sapi
24	507.2	15.1	595	6	CQ100724	CQ100724 Sequence
25	507.2	15.1	595	6	CQ139716	CQ139716 Sequence
26	507.2	15.1	595	6	CQ261044	CQ261044 Sequence
27	507.2	15.1	595	6	CQ335189	CQ335189 Sequence
28	455.2	13.5	499	6	CQ113866	CQ113866 Sequence
29	455.2	13.5	499	6	CQ152744	CQ152744 Sequence
30	455.2	13.5	499	6	CQ273644	CQ273644 Sequence
31	455.2	13.5	499	6	CQ347943	CQ347943 Sequence
32	440.8	13.1	444	6	CQ729107	CQ729107 Sequence
33	384	11.4	923	10	BC019605	BC019605 Mus muscu
34	376.4	11.2	458	6	BD269355	BD269355 Gene show
35	371.2	11.0	392	6	AX071769	AX071769 Sequence
36	365.8	10.9	374	6	BD071774	BD071774 Secreted
37	356.8	10.6	387	6	AX071783	AX071783 Sequence
38	336	10.0	3458	5	BC073528	BC073528 Xenopus 1
39	331.4	9.8	341	6	CQ431059	CQ431059 Sequence
40	330.4	9.8	340	6	CQ422192	CQ422192 Sequence
41	303	9.0	163681	9	AC007750	AC007750 Homo sapi
42	285.4	8.5	301	6	I50897	I50897 Sequence 7
43	285.4	8.5	301	6	AR494597	AR494597 Sequence
44	285.4	8.5	301	6	AX806273	AX806273 Sequence
45	284.4	8.5	223513	2	AC111919	AC111919 Rattus no

ALIGNMENTS

RESULT 1	CQ776470	3380 bp	DNA	linear	PAT 11-MAR-2004
LOCUS	CQ776470	Sequence 156 from Patent EP1394274.			
DEFINITION	CQ776470				
ACCESSION	CQ776470.1	GI:45379860			
VERSION					
KEYWORDS					
SOURCE					
ORGANISM	Homo sapiens (human)				
REFERENCE	1	Ohtani,N., Sugita,Y., Yamaya,M., Kubo,H., Nagai,H. and Izuhara,K.			
AUTHORS					
TITLE		Methods of testing for bronchial asthma or chronic obstructive,			
JOURNAL		pulmonary disease			
FEATURES		Patent: EP 1394274-A 156 03-MAR-2004;			
source		Genox Research, Inc. (JP)			
ORIGIN		Location/Qualifiers			
		1..3380			
		/organism="Homo sapiens"			
		/mol_type="unassigned DNA"			
		/db_xref="taxon:9606"			
Query Match	100.0%;	Score 3365;	DB 6;	Length 3380;	
Best Local Similarity	100.0%;	Pred. No. 0;			
Matches 3365;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
QY	1	GGCGCGCGGCTGAGAGCCCTGTGACAACTCGTCATGTCAGGACAGAGCCGGTGAAC	60		
DB	1	GCGCGCGGCTGAGAGCCCTGTGACAACTCGTCATGTCAGGACAGAGCCGGTGAAC	60		
QY	61	CCTGCTTCTTAAGTGGGACGGACAGCGGACACATTTCACTGTCCCGACAGCAA	120		
DB	61	CCTGCTTCTTAAGTGGGACGGACAGCGGACACATTTCACTGTCCCGACAGCAA	120		
QY	121	CAGCACCATCTGCTTGGAGAACCTCTCCCTTCTCTGAGAAAGATGCGAATGGG	180		
DB	121	CAGCACCATCTGCTTGGAGAACCTCTCCCTTCTCTGAGAAAGATGCGAATGGG	180		
QY	181	TATTCACAGACGAGATTTCCGCTATCTCATCTGTCGTTAGGGCCAGGATGAATG	240		

QY 2401 GCTGAAGTAGAGTCAAAGCCCAACCATCTGATTGGAGCTGGACACAGCAGTGAAGTTCAAA 2460
Db |||||||
2401 GCTGAAGTAGAGTCAAAGCCCAACCATCTGATTGGAGCTGGACACAGCAGTGAAGTTCAAA 2460
QY 2461 CCCATGACACAGAATGAACAAAAAGAGTCAATAGTAGTAATTTGCGACTGGAAAAATCAAT 2520
Db |||||||
2461 CCCATGACACAGAATGAACAAAAAGAGTCAATAGTAGTAATTTGCGACTGGAAAAATCAAT 2520
QY 2521 CTGCTTATCGTACCACAGTGGCAGAAGAGTCTGGATATTAAGAAATGTAACATTTGTT 2580
Db |||||||
2521 CTGCTTATCGTACCACAGTGGCAGAAGAGTCTGGATATTAAGAAATGTAACATTTGTT 2580
QY 2581 ATCCGTTATGTTCTGTCACCAATGAATAGCCATGTTCCAGGCCCGTGGTCGAGCCAGA 2640
Db |||||||
2581 ATCCGTTATGTTCTGTCACCAATGAATAGCCATGTTCCAGGCCCGTGGTCGAGCCAGA 2640
QY 2641 GCTGATGAGACCACTACGTCCTGTTGCTCACAGTGGTTCAGGAGTTATCGAACATGAG 2700
Db |||||||
2641 GCTGATGAGACCACTACGTCCTGTTGCTCACAGTGGTTCAGGAGTTATCGAACATGAG 2700
QY 2701 ACAGTTAATGATTTCCGAGAGAAGATGATGTATAAAGCTATACATTTGTTTCAAAATATG 2760
Db |||||||
2701 ACAGTTAATGATTTCCGAGAGAAGATGATGTATAAAGCTATACATTTGTTTCAAAATATG 2760
QY 2761 AAACGAGAGGAGTATGCTCATAGATTTTGGAAATTACAGATGCAAAAGTATAATGAAAAAG 2820
Db |||||||
2761 AAACGAGAGGAGTATGCTCATAGATTTTGGAAATTACAGATGCAAAAGTATAATGAAAAAG 2820
QY 2821 AAAATGAAACCAAGAGAAATATTGCCAAGCATTACAAGATAACCCATCACTAATAACT 2880
Db |||||||
2821 AAAATGAAACCAAGAGAAATATTGCCAAGCATTACAAGATAACCCATCACTAATAACT 2880
QY 2881 TTCCTTTGCAAAAACTGCAGTGTGCTAGCCCTGTTCTGGGAAGATATCCATGTAATTGAG 2940
Db |||||||
2881 TTCCTTTGCAAAAACTGCAGTGTGCTAGCCCTGTTCTGGGAAGATATCCATGTAATTGAG 2940
QY 2941 AAAATGCAATCACGTCAATATGACCCAGAAATTCAGGAATTTACATTTGTAAGAGAAAAAC 3000
Db |||||||
2941 AAAATGCAATCACGTCAATATGACCCAGAAATTCAGGAATTTACATTTGTAAGAGAAAAAC 3000
QY 3001 AAAGCACTGCAAAAAGAGTGTGCCGACTATCAAAATAAATGGTGAATCATCTGCAAAATGT 3060
Db |||||||
3001 AAAGCACTGCAAAAAGAGTGTGCCGACTATCAAAATAAATGGTGAATCATCTGCAAAATGT 3060
QY 3061 GGCCAGGCTTGGGGAACAATGATGTGTCACAAAGGCTTAGATTGCTTCTCAAAATA 3120
Db |||||||
3061 GGCCAGGCTTGGGGAACAATGATGTGTCACAAAGGCTTAGATTGCTTCTCAAAATA 3120
QY 3121 AGGAATTTTGTAGTGGTTTTCAAAAATAATTCAACAAAGAAACAATACAAAAAGTGGGTA 3180
Db |||||||
3121 AGGAATTTTGTAGTGGTTTTCAAAAATAATTCAACAAAGAAACAATACAAAAAGTGGGTA 3180
QY 3181 GAATTACCTATCACATTTCCCAATCTTGACTATTTCAGAAATGCTGTTTATTTAGTGATGAG 3240
Db |||||||
3181 GAATTACCTATCACATTTCCCAATCTTGACTATTTCAGAAATGCTGTTTATTTAGTGATGAG 3240
QY 3241 GATTAGCACTTGATTGAAGATTTCTTTAAAAATACTATCAGTTAAACATTTAATATGATTA 3300
Db |||||||
3241 GATTAGCACTTGATTGAAGATTTCTTTAAAAATACTATCAGTTAAACATTTAATATGATTA 3300
QY 3301 TGATTAATGATTCATTATGCTACAGAACTGACATAAGAAATCAATAAAATGATTGTTTAA 3360
Db |||||||
3301 TGATTAATGATTCATTATGCTACAGAACTGACATAAGAAATCAATAAAATGATTGTTTAA 3360
QY 3361 CTCTG 3365
Db |||||
3361 CTCTG 3365

RESULT 2
AF095844
LOCUS AF095844 3380 bp mRNA linear PRI 23-JAN-2002
DEFINITION Homo sapiens melanoma differentiation associated protein-5 (MDA5)

ACCESSION mRNA, complete cds.
VERSION AF095844
KEYWORDS AF095844.1 GI:11344593
SOURCE
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3380)
AUTHORS Kang,D.C., Gopalkrishnan,R.V., Wu,Q., Jankowsky,E., Pyle,A.M. and Fisher,P.B.
TITLE mda-5: An interferon-inducible putative RNA helicase with double-stranded RNA-dependent ATPase activity and melanoma growth-suppressive properties
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (2), 637-642 (2002)
MEDLINE 21664412
PUBMED 11805321
REFERENCE 2 (bases 1 to 3380)
AUTHORS Kang,D.-C. and Fisher,P.B.
TITLE Direct Submission
JOURNAL Submitted (30-SEP-1998) Urology, Columbia University, P&S, 630 W168th, New York, NY 10032, USA
FEATURES
Location/Qualifiers
1..3380
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_line="HO-1"
/tissue_type="melanoma"
1..3380
/gene="MDA5"
169..3246
/gene="MDA5"
/codon_start=1
/product="melanoma differentiation associated protein-5"
/protein_id="AAG34368.1"
/db_xref="GI:11344594"
/translation="MSGNGYSTDENFRYLISCFRARVVMYIQVEPVLDTLTFLEAEVKE
QIORTVATSGNMOAVELLSTLEKGVHLGWTREFVEALRRTGSPLAARYMNPDLTDL
PSPFENAHDEYLQNLNLLQPTLVLDKLVDRDVKCMEELLTIEDNRIRAAENNGN
ESGVRELLKRIVQENWFSFLNVLQTNNELVQELTSGDCSESNABENLSQVDPG
QVEEQLLSTTVQNLKVBWGMENNSSESPADSSVSESDTSLAEGSVSCLDSESLGH
NSNMGSDSGTMSGSDSEENVAARASPEPELQLRPYQMEVAQPALEGKNIIICLPTSGS
KTRVAVYIAKDLDDKKKASEPGKVIVLVNKVLLVEQLFRKEFPFLKKWYRVIGLSG
DTQLKISFPEVVKSCDIIISTAQILENSLNLNENGEDAGVQLSDFSLIIDECHHTNK
EAVYNNIMRHYLMQKLKNNRLKKENKPVIPQLIGLTASPGVGGATQAKAEHILK
LCANLDAFTIKTVKENLDQKNQIQEPCKFAIADATREDPFKEKLLLEIMTRIQTQYCO
MSPMSDFGTPQYEQWAIQMEKKAAGKGRKERVCAEHLRKYNEALQINDTIRMIDAYT
HLETFYNEEKDKKFAVEDSDSDEGGDDCYDGEDDEDLKKPLKLDDETRFLMTLFFE
NNKMLKRLAENPEYENEKLTCLRNTIMEQYTRTEESARGIIFTKTQSAYALSQWITE
NEKFAEVGVKAHLIGAGHSSEFKPMTONEQKEVISKFRGTGKINLLIATTVAEEGLDI
KECNIVIRYGLVTNEIAMVQARGARADESTYVLVAHSGSGVIEHETVNDFREKMYK
AHCQVQNMKPEEYAHKILELQMQSIMKMKTKRNIAGHYKNNPSLITFLCKNCVSLA
CSGEDIHVIEKMHVNMTPFEKELYIVRENKALQKKCADYQINGEIIICKGQAWGTM
VHKGDLPLCLKIRNFVVFVKNNSTKKQYKKWVELPITFPNLDYSECCLFSDDED"

ORIGIN

Query Match 100.0%; Score 3365; DB 9; Length 3380;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3365; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCGCGCGCGCCTGAGAGCCCTGTGGACAACCTCGTCAATTGTGAGGCACAGCGGTAGAC 60
Db |||||||
1 GCGCGCGCGCCTGAGAGCCCTGTGGACAACCTCGTCAATTGTGAGGCACAGCGGTAGAC 60
QY 61 CCTGCTTCTCTAAGTGGCAGCGGACAGCGGCACGACATTTACCTGTCTCCCGACACAA 120
Db |||||||
61 CCTGCTTCTCTAAGTGGCAGCGGACAGCGGCACGACATTTACCTGTCTCCCGACACAA 120
QY 121 CAGCACCATCTGCTTGGGAGAACCTCTCCCTTCTCTGAGAAAGAAAGATGTCGAATGGG 180
Db |||||||
121 CAGCACCATCTGCTTGGGAGAACCTCTCCCTTCTCTGAGAAAGAAAGATGTCGAATGGG 180

QY	181	TATTCCACAGACGAGAATTTCCGCTATCTCATCTCGTGTCTTACAGGCCAGGGTGAAAAATG	240
Db	181		240
QY	241	TATTCCACAGACGAGAATTTCCGCTATCTCATCTCGTGTCTTACAGGCCAGGGTGAAAAATG	300
Db	241		300
QY	301	TACATCCAGGTGGAGCCTGTGCTGGACTACCTGACCTTTCTGCTCCTGCAGAGGTGAAGGAG	360
Db	301		360
QY	361	TACATCCAGGTGGAGCCTGTGCTGGACTACCTGACCTTTCTGCTCCTGCAGAGGTGAAGGAG	420
Db	361		420
QY	421	ACCTTGGAGAAAGGAGTCTGGCACCTTGGTTGGACTCGGGAATTCGTGGAGGCCCTCCGG	480
Db	421		480
QY	481	AGAACCGGCAGCCCTCTGGCCGCCCGCTACATGAACCTTGAGCTCACGGACTTGCCCTCT	540
Db	481		540
QY	541	AGAACCGGCAGCCCTCTGGCCGCCCGCTACATGAACCTTGAGCTCACGGACTTGCCCTCT	600
Db	541		600
QY	601	CCATCGTTTGAGAACGCTCATGATGAATATCTCCAACCTGCTGAACCTCCTTCAGGCCCACT	660
Db	601		660
QY	661	CCATCGTTTGAGAACGCTCATGATGAATATCTCCAACCTGCTGAACCTCCTTCAGGCCCACT	720
Db	661		720
QY	721	CTGGTGGACAAAGCTTCTAGTTAGAGACGCTCTTGGATAAGTGCAATGGAGGAGGAACTGTTG	780
Db	721		780
QY	781	CTGGTGGACAAAGCTTCTAGTTAGAGACGCTCTTGGATAAGTGCAATGGAGGAGGAACTGTTG	840
Db	781		840
QY	841	CTGGTGGACAAAGCTTCTAGTTAGAGACGCTCTTGGATAAGTGCAATGGAGGAGGAACTGTTG	900
Db	841		900
QY	901	CTGGTGGACAAAGCTTCTAGTTAGAGACGCTCTTGGATAAGTGCAATGGAGGAGGAACTGTTG	960
Db	901		960
QY	961	CTGGTGGACAAAGCTTCTAGTTAGAGACGCTCTTGGATAAGTGCAATGGAGGAGGAACTGTTG	1020
Db	961		1020
QY	1021	CTGGTGGACAAAGCTTCTAGTTAGAGACGCTCTTGGATAAGTGCAATGGAGGAGGAACTGTTG	1080
Db	1021		1080
QY	1081	CTGGTGGACAAAGCTTCTAGTTAGAGACGCTCTTGGATAAGTGCAATGGAGGAGGAACTGTTG	1140
Db	1081		1140
QY	1141	CTGGTGGACAAAGCTTCTAGTTAGAGACGCTCTTGGATAAGTGCAATGGAGGAGGAACTGTTG	1200
Db	1141		1200
QY	1201	CTGGTGGACAAAGCTTCTAGTTAGAGACGCTCTTGGATAAGTGCAATGGAGGAGGAACTGTTG	1260
Db	1201		1260
QY	1261	CTGGTGGACAAAGCTTCTAGTTAGAGACGCTCTTGGATAAGTGCAATGGAGGAGGAACTGTTG	1320
Db	1261		1320

Db	1261	 AAGTACTGCTAGTTGAACAGCTCTTCCGCAAGGAGTTCCAACCACTTTTGAAGAAATGG	1320
Qy	1321	 TATCGTGTATTGGATTAAAGTGGTGATACCCAACTGAAAAATATCATTTCCAGAAAGTTGTC	1380
Db	1321	 TATCGTGTATTGGATTAAAGTGGTGATACCCAACTGAAAAATATCATTTCCAGAAAGTTGTC	1380
Qy	1381	 AAGTCCGTGTGATATTATATCAGTACAGTCAAATCCTTGAAAACTCCCTCTTTAAACCTTG	1440
Db	1381	 AAGTCCGTGTGATATTATATCAGTACAGTCAAATCCTTGAAAACTCCCTCTTTAAACCTTG	1440
Qy	1441	 GAAAAATGGAGAGATGCTGGTGTTCATTTGTTCAGACTTTTCCCTCATTTATCATTTGATGAA	1500
Db	1441	 GAAAAATGGAGAGATGCTGGTGTTCATTTGTTCAGACTTTTCCCTCATTTATCATTTGATGAA	1500
Qy	1501	 TGTCATCACACCAACAAAGACAGTGTATAATAACATCATGAGGCATTATTGATGCAG	1560
Db	1501	 TGTCATCACACCAACAAAGACAGTGTATAATAACATCATGAGGCATTATTGATGCAG	1560
Qy	1561	 AAGTTGAAAAACAATAGACTCAAGAAAGAAAAACAAACCAAGTGATTCCTTCAGATA	1620
Db	1561	 AAGTTGAAAAACAATAGACTCAAGAAAGAAAAACAAACCAAGTGATTCCTTCAGATA	1620
Qy	1621	 CTGGGACTAACAGCTTACCTGGTGTGGAGGGGCCACGAAGCAAGCCAAAGCTGAAGAA	1680
Db	1621	 CTGGGACTAACAGCTTACCTGGTGTGGAGGGGCCACGAAGCAAGCCAAAGCTGAAGAA	1680
Qy	1681	 CACATTTTAAAACTATGTGCCAATCTTGATGCAATTTACTATTAACCTGTTAAAGAAAAC	1740
Db	1681	 CACATTTTAAAACTATGTGCCAATCTTGATGCAATTTACTATTAACCTGTTAAAGAAAAC	1740
Qy	1741	 CTTGATCAACTGAAAAACCAATACAGGAGCCATGCAAGAGTTTGCCATTCAGATGCA	1800
Db	1741	 CTTGATCAACTGAAAAACCAATACAGGAGCCATGCAAGAGTTTGCCATTCAGATGCA	1800
Qy	1801	 ACCAGAGAAGATCCATTTAAAGAGAAAACCTTCTAGAAAATAATGACAAGGATTCAAACTTAT	1860
Db	1801	 ACCAGAGAAGATCCATTTAAAGAGAAAACCTTCTAGAAAATAATGACAAGGATTCAAACTTAT	1860
Qy	1861	 TGTCAAATGAGTCCAATGTCCAGATTTTGGAACTCAACCCCTATGAACAATGGGCCATTCAA	1920
Db	1861	 TGTCAAATGAGTCCAATGTCCAGATTTTGGAACTCAACCCCTATGAACAATGGGCCATTCAA	1920
Qy	1921	 ATGGAAAAAAAAGCTGCAAAAAAAGGAAATCGCAAGAACGTGTTTGTGCGAGAACATTTG	1980
Db	1921	 ATGGAAAAAAAAGCTGCAAAAAAAGGAAATCGCAAGAACGTGTTTGTGCGAGAACATTTG	1980
Qy	1981	 AGGAAGTACAATGAGGCCCTACAAATTAATGACACAATTCGAATGATAGATGCGTATACT	2040
Db	1981	 AGGAAGTACAATGAGGCCCTACAAATTAATGACACAATTCGAATGATAGATGCGTATACT	2040
Qy	2041	 CATCTTGAAACCTTTCTATAATGAAGAGAAAGATAAGAAAGTTTGCAGTCATAGAAGATGAT	2100
Db	2041	 CATCTTGAAACCTTTCTATAATGAAGAGAAAGATAAGAAAGTTTGCAGTCATAGAAGATGAT	2100
Qy	2101	 AGTGATGAGGGTGGTGATGATGAGTATTGTGTGATGGTGATGAAGATGAGGATGATTTAAAG	2160
Db	2101	 AGTGATGAGGGTGGTGATGATGAGTATTGTGTGATGGTGATGAAGATGAGGATGATTTAAAG	2160
Qy	2161	 AAACCTTTGAAACTGGATGAAACAGATAGATTTCTCATGACTTTATTTTTTGAACCAAT	2220
Db	2161	 AAACCTTTGAAACTGGATGAAACAGATAGATTTCTCATGACTTTATTTTTTGAACCAAT	2220
Qy	2221	 AAAAATGTTGAAAAGGCTGGCTGAAAAACCCAGATAATGAAAAATGAAAAAGCTGACCAATTA	2280
Db	2221	 AAAAATGTTGAAAAGGCTGGCTGAAAAACCCAGATAATGAAAAATGAAAAAGCTGACCAATTA	2280
Qy	2281	 AGAAATACCATAATGGAGCAATATACTAGGACTGAGGAATCAGCACGAGGAATAATCTTT	2340
Db	2281	 AGAAATACCATAATGGAGCAATATACTAGGACTGAGGAATCAGCACGAGGAATAATCTTT	2340
Qy	2341	 ACAAAAACACGACAGTGCATATGCGCTTTCCCACTGGATTACTGAAAAATGAAAAATTT	2400
Db	2341	 ACAAAAACACGACAGTGCATATGCGCTTTCCCACTGGATTACTGAAAAATGAAAAATTT	2400

Db	2341	ACAAAAACACGACAGAGTGCAATATGGCGCTTCCAGTGGATTACTGAAAAATGAAAAATTT	2400
QY	2401	GCTGAAGTAGGAGTCAAAGCCCCACCATCTGATTGGAGCTGGACAGCAGTGAGTTCAAA	2460
Db	2401	GCTGAAGTAGGAGTCAAAGCCCCACCATCTGATTGGAGCTGGACAGCAGTGAGTTCAAA	2460
QY	2461	CCCATGACACAGAAATGAACAAAAAGAGTCATTAGTAGTAATTTCCGACTGGAAAAATCAAT	2520
Db	2461	CCCATGACACAGAAATGAACAAAAAGAGTCATTAGTAGTAATTTCCGACTGGAAAAATCAAT	2520
QY	2521	CTGCTTATCGCTACCACAGTGGCAGAGAGAGGTCTGGATATTAAGAATGTAACATTGTT	2580
Db	2521	CTGCTTATCGCTACCACAGTGGCAGAGAGAGGTCTGGATATTAAGAATGTAACATTGTT	2580
QY	2581	ATCCGTTATGGTCTCGTCACCAATGAATAGCCATGGTCCAGGCCCGTGGTCGAGCCAGA	2640
Db	2581	ATCCGTTATGGTCTCGTCACCAATGAATAGCCATGGTCCAGGCCCGTGGTCGAGCCAGA	2640
QY	2641	GCTGATGAGAGCACCTACGTCCTGGTTGCTCACAGTGGTTCAGGAGTTATCGAACATGAG	2700
Db	2641	GCTGATGAGAGCACCTACGTCCTGGTTGCTCACAGTGGTTCAGGAGTTATCGAACATGAG	2700
QY	2701	ACAGTTAATGATTTCCGAGAGAAAGATGATGTATAAAGCTATACATTGTTCAAAAATATG	2760
Db	2701	ACAGTTAATGATTTCCGAGAGAAAGATGATGTATAAAGCTATACATTGTTCAAAAATATG	2760
QY	2761	AAACCAGAGGAGTATGCTCATAAGATTTTGGAAATTACAGATGCCAAAGTATAATGGAAAG	2820
Db	2761	AAACCAGAGGAGTATGCTCATAAGATTTTGGAAATTACAGATGCCAAAGTATAATGGAAAG	2820
QY	2821	AAAAATGAACCAAGAGAAATATTGCCAAGCATTACAAGAATAACCCATCACTAATAACT	2880
Db	2821	AAAAATGAACCAAGAGAAATATTGCCAAGCATTACAAGAATAACCCATCACTAATAACT	2880
QY	2881	TTCCCTTTGCAAAAACCTGCAGTGTGCTAGCCTGTTCTGGGGAAGATATCCATGTAATTGAG	2940
Db	2881	TTCCCTTTGCAAAAACCTGCAGTGTGCTAGCCTGTTCTGGGGAAGATATCCATGTAATTGAG	2940
QY	2941	AAAATGCATCAGTCAATATGACCCCAGAAATCAAGGAACTTTACATTGTAAGAGAAAC	3000
Db	2941	AAAATGCATCAGTCAATATGACCCCAGAAATCAAGGAACTTTACATTGTAAGAGAAAC	3000
QY	3001	AAAGCACTGCAAAAAGAGTGTGCCGACTATCAAAATAAATGGTGAATCATCTGCAAAATGT	3060
Db	3001	AAAGCACTGCAAAAAGAGTGTGCCGACTATCAAAATAAATGGTGAATCATCTGCAAAATGT	3060
QY	3061	GGCCAGGCTTGGGGAACAATGATGGTGCAAAAAGGCTTAGATTTCGCTTGCTCAAAATA	3120
Db	3061	GGCCAGGCTTGGGGAACAATGATGGTGCAAAAAGGCTTAGATTTCGCTTGCTCAAAATA	3120
QY	3121	AGGAATTTTGTAGTGGTTTTCAAAAATAATTCAACAAAGAAACAATACAAAAAGTGGGTA	3180
Db	3121	AGGAATTTTGTAGTGGTTTTCAAAAATAATTCAACAAAGAAACAATACAAAAAGTGGGTA	3180
QY	3181	GAATTACCTATCACATTTCCCAATCTTGACTATTTCAGAAATGCTGTTATTAGTGATGAG	3240
Db	3181	GAATTACCTATCACATTTCCCAATCTTGACTATTTCAGAAATGCTGTTATTAGTGATGAG	3240
QY	3241	GATTAGCACTTGATTGAAGATTCTTTTAAATACTATCAGTTAAACATTTAATATGATTA	3300
Db	3241	GATTAGCACTTGATTGAAGATTCTTTTAAATACTATCAGTTAAACATTTAATATGATTA	3300
QY	3301	TGATTAATGATTCAATTATGCTACAGAACTGACATAAGAATCAATAAAATGATTGTTTTA	3360
Db	3301	TGATTAATGATTCAATTATGCTACAGAACTGACATAAGAATCAATAAAATGATTGTTTTA	3360
QY	3361	CTCTG 3365	
Db	3361	CTCTG 3365	

LOCUS	AX300832	3372 bp	DNA	linear	PAT 30-NOV-2001
DEFINITION	Sequence 1 from Patent WO0185955.				
ACCESSION	AX300832				
VERSION	AX300832.1 GI:17382110				
KEYWORDS					
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1				
AUTHORS	Bahr, G., Cocude, C. and Capron, A.				
TITLE	Rh116 polypeptides and its fragments and polynucleotides encoding said polypeptides and therapeutic uses				
JOURNAL	Patent: WO 0185955-A 1 15-NOV-2001;				
	Istac (FR); INSTITUT PASTEUR DE LILLE (FR)				
FEATURES	Location/Qualifiers				
source	1. .3372				
	/organism="Homo sapiens"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:9606"				
CDS	155. .3232				
	/note="unnamed protein product"				
	/codon_start=1				
	/protein_id="CAD13049.1"				
	/db_xref="GI:17382111"				
	/translation="MSGYSTDENFRYLISCFRARKVMYIQVEPVLDTFLPAEVKEQIORTVATSGNMQAVELLSTLEKGVHLGWTREFVEALRRTGSPLAARYNMPELTDLPSPFENAHDEYQLLNLLQPTLVLDKLLVRDLDKMEELLTIEDNRNRIAAENNGNBSGVRELLKRIVQKENWFSAPLNLVRLQTNNELVQLTSGDCSESNAEIENLSQVDGPQVEEQLLSTTVQPNLEKEVGMENSSSFADSVSVESDTSLAEGSVSCLDSESLGHNSNMGSDSGTMGSDSDENVAARASPEPELQLRPYQMEVAQPALEKKNIIICLPTGSGKTRVAVYIAKDLHDKKASEPGKIVLVNKLVLVEQLFRKEFQFLKKNYRVIGLSGDTQLKISFPEVVKSDIISTAQILENSLLNLENGEDAGVQLSDFSFIIDCHHTNKEAVYNIMRHYLMQKLNRLKKNKPVIPLPQILGLTASPGVGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQIQEPCKFAIADATREDPFKEKLEIMTRIQTYYCOMSPMDFGTQPYEQWAIQMEKKAKEGNRKESVCAEHLRKYNKALQINDTIRMIDAYTHLETFYNEEKDKKFAVIDDDSDDEGDDDEYCDGDEDDDLKKPLKLDDETRFLMTLFFENNKMLKRLAENPEYENEKLTKLNTIMEQYTRTEESARGIIFTKTRQSAVALSQWITENEKPAEVGVKAHLIGAHSGSEFPKPMTQNEQKVISKFTGKINLLIATTVAEGLDIKECNIVIRYGLVTNEIAMVQARGARADESTYVLVHSGSGVIERETVNDFREKMYKAIHCQVQNMKPEEYAHKILELQMQSIMKMKTKRNIASHKYNPSLITFLCKNCSVLA CSGEDIHVIEKMHVNMTPPEFKELYIVRENKALOKKADYQINGEIIICKGQAWGTMVHKGLDLPCLKIRNFVVVPKNNSTKKQYKKWVLPITFPNLDYSECCLFSDSD"				
ORIGIN					
Query Match	99.2%;	Score 3338.8;	DB 6;	Length 3372;	
Best Local Similarity	99.8%;	Pred. No. 0;			
Matches 3343;	Conservative 0;	Mismatches 7;	Indels 0;	Gaps 0;	
QY	15	GAGCCCTGTGGACAAACCTCGTCAATTGTTCAGGCACAGAGCGGTAGACCCTGCTTCTCTAAG	74		
Db	1	GGGCCCTGTGGACAAACCTCGTCAATTGTTCAGGCACAGAGCGGTAGACCCTGCTTCTCTAAG	60		
QY	75	TGGGCAGCGGACAGCGGCACGCACATTTACCTGTCCCAGACACAACAGCACCATCTGCT	134		
Db	61	TGGGCAGCGGACAGCGGCACGCACATTTACCTGTCCCAGACACAACAGCACCATCTGCT	120		
QY	135	TGGGAGAACCTCTCCCTTCTCTGAGAAAGAAAGATGTCGAATGGGTATTCACACAGACA	194		
Db	121	TGGGAGAACCTCTCCCTTCTCTGAGAAAGAAAGATGTCGAATGGGTATTCACACAGACA	180		
QY	195	GAATTTCCGCTATCTCATCTCGTCTTCAGGCCAGGGTGAAAAATGTACATCCAGGTGGA	254		
Db	181	GAATTTCCGCTATCTCATCTCGTCTTCAGGCCAGGGTGAAAAATGTACATCCAGGTGGA	240		
QY	255	GCCTGTCTGGACTACCTGACCTTTCTGCCTGCAGAGGTGAAGGAGCAGATTTCAGAGGAC	314		
Db	241	GCCTGTCTGGACTACCTGACCTTTCTGCCTGCAGAGGTGAAGGAGCAGATTTCAGAGGAC	300		
QY	315	AGTCGCCACCTCCGGGAACATGCAGGCAGTTGAACCTGCTGCTGAGCACCTTGGAGAGGG	374		
Db	301	AGTCGCCACCTCCGGGAACATGCAGGCAGTTGAACCTGCTGCTGAGCACCTTGGAGAGGG	360		

QY 375 AGTCTGGCACCTTGGTTGGACTCGGGAATTCGTGAGGCCCTCCGGAGAACCCGGCAGCCC 434
Db 361 AGTCTGGCACCTTGGTTGGACTCGGGAATTCGTGAGGCCCTCCGGAGAACCCGGCAGCCC 420
QY 435 TCTGGCCGCCGCTACATGAACCTGAGCTCACGGACTTGCCTCTCCATCGTTTGAGAA 494
Db 421 TCTGGCCGCCGCTACATGAACCTGAGCTCACGGACTTGCCTCTCCATCGTTTGAGAA 480
QY 495 CGCTCATGATGAATATCTCCAACTGCTGAACCTCCTTCAGCCCACTCTGTGGACAAGCT 554
Db 481 CGCTCATGATGAATATCTCCAACTGCTGAACCTCCTTCAGCCCACTCTGTGGACAAGCT 540
QY 555 TCTAGTTAGACAGCTCTTGGATTAAGTGCATGGAGGAGAACTGTTGACAATTAAGACAG 614
Db 541 TCTAGTTAGACAGCTCTTGGATTAAGTGCATGGAGGAGAACTGTTGACAATTAAGACAG 600
QY 615 AAACCGGATTGCTGCTGCAGAAAAAACAATGGAATCAGGTGTAAGAGAGCTACTAAA 674
Db 601 AAACCGGATTGCTGCTGCAGAAAAAACAATGGAATCAGGTGTAAGAGAGCTACTAAA 660
QY 675 AAGGATTGTGAGAAAAAGAAAACCTGGTTCTCTGCAATTTCTGAATGTTCTTCGTCAAACAGG 734
Db 661 AAGGATTGTGAGAAAAAGAAAACCTGGTTCTCTGCAATTTCTGAATGTTCTTCGTCAAACAGG 720
QY 735 AAACAAATGAATGTTCCAAGAGTTAAACAGGCTCTGATTGCTCAGAAAGCAATCATCAGAAATCATCTTT 794
Db 721 AAACAAATGAATGTTCCAAGAGTTAAACAGGCTCTGATTGCTCAGAAAGCAATCATCAGAAATCATCTTT 780
QY 795 TGAGAAATTTATCAAAAGTTGATGGTCTCAAGTGGAAGCAACTTCTTCAACCCACAGT 854
Db 781 TGAGAAATTTATCAAAAGTTGATGGTCTCAAGTGGAAGCAACTTCTTCAACCCACAGT 840
QY 855 TCAGCCAAATCTGGAGAAGGAGGCTCTGGGCGCATGGAGAAATTAATCATCAGAAATCATCTTT 914
Db 841 TCAGCCAAATCTGGAGAAGGAGGCTCTGGGCGCATGGAGAAATTAATCATCAGAAATCATCTTT 900
QY 915 TGCAGATTCTTCTGTAGTTTCAGAAATCAGACACAAGTTTGGCAGAAAGGAAGTGTCAAGCTG 974
Db 901 TGCAGATTCTTCTGTAGTTTCAGAAATCAGACACAAGTTTGGCAGAAAGGAAGTGTCAAGCTG 960
QY 975 CTTAGATGAAAGTCTTGGACATAACAGCAACATGGGCAGTGATTCAGGCACCATGGGAAG 1034
Db 961 CTTAGATGAAAGTCTTGGACATAACAGCAACATGGGCAGTGATTCAGGCACCATGGGAAG 1020
QY 1035 TGATTTCAGATGAAGAGAAATGTGGCAGCAAGAGCATCCCCGGAGCCAGAACTCCAGCTCAG 1094
Db 1021 TGATTTCAGATGAAGAGAAATGTGGCAGCAAGAGCATCCCCGGAGCCAGAACTCCAGCTCAG 1080
QY 1095 GCCTTACCAAATGGAAAGTTGCCAGCCAGCCCTTGGAAAGGAAGAAATATCATCATCTGCCT 1154
Db 1081 GCCTTACCAAATGGAAAGTTGCCAGCCAGCCCTTGGAAAGGAAGAAATATCATCATCTGCCT 1140
QY 1155 CCTTACAGGGAGTGGAAAAAACCCAGAGTGGCTGTTTACATTGCCAAGGATCACTTAGACAA 1214
Db 1141 CCTTACAGGGAGTGGAAAAAACCCAGAGTGGCTGTTTACATTGCCAAGGATCACTTAGACAA 1200
QY 1215 GAAGAAAAAAGCATCTGAGCTGGAAAAAGTTATAGTTCTTGTCAAAGGTACTGCTAGT 1274
Db 1201 GAAGAAAAAAGCATCTGAGCTGGAAAAAGTTATAGTTCTTGTCAAAGGTACTGCTAGT 1260
QY 1275 TGAACAGCTCTTCCGCAAGGAGTTCCAACCAATTTTGAAGAAATGGTATCGTGTATTGG 1334
Db 1261 TGAACAGCTCTTCCGCAAGGAGTTCCAACCAATTTTGAAGAAATGGTATCGTGTATTGG 1320
QY 1335 ATTAAGTGGTGATACCCCACTGAAAAATATCATTTCCAGAAAGTTGTCAAGTCTGTGATAT 1394
Db 1321 ATTAAGTGGTGATACCCCACTGAAAAATATCATTTCCAGAAAGTTGTCAAGTCTGTGATAT 1380
QY 1395 TATTATCAGTACAGCTCAAAATCCTTGAAAACTCCCTCTTAAACTTGGAAAAATGGAGAAGA 1454
Db 1381 TATTATCAGTACAGCTCAAAATCCTTGAAAACTCCCTCTTAAACTTGGAAAAATGGAGAAGA 1440
QY 1455 TGCTGGTGTTCAAATTGTCAGACTTTTCCCTCATTTATCATGTGATGAATGTCATCACACCAA 1514

Db 1441 TGCTGGTGTCAAATGTGCAAGCTTTTCTTCAATTATCATGTGATGAATGTCATCACACCAA 1500
QY 1515 CAAAGAAAGCAGTGTATAATAACATCATGAGGCATTAATTGATGCAGAAAGTTGAAAAACAA 1574
Db 1501 CAAAGAAAGCAGTGTATAATAACATCATGAGGCATTAATTGATGCAGAAAGTTGAAAAACAA 1560
QY 1575 TAGACTCAAGAAAGAAACAAACCAGTGATTTCCCTTCCCTCAGATACCTGGGACTAACAGC 1634
Db 1561 TAGACTCAAGAAAGAAACAAACCAGTGATTTCCCTTCCCTCAGATACCTGGGACTAACAGC 1620
QY 1635 TTCACTCTGGTGTGGAGGGGCCACGAAGCAAGCCAAAGCTGAAGAACACACATTTTAAAACT 1694
Db 1621 TTCACTCTGGTGTGGAGGGGCCACGAAGCAAGCCAAAGCTGAAGAACACACATTTTAAAACT 1680
QY 1695 ATGTGCCAATCTTGATGCTATTTACTATTAAACCTGTTAAAGAAAAACCTTGATCAACTGAA 1754
Db 1681 ATGTGCCAATCTTGATGCTATTTACTATTAAACCTGTTAAAGAAAAACCTTGATCAACTGAA 1740
QY 1755 AAACCAAATACAGGAGCCATGCAAGAAAGTTTGCCATTGCGAGATGCAACCCAGAGAGATCC 1814
Db 1741 AAACCAAATACAGGAGCCATGCAAGAAAGTTTGCCATTGCGAGATGCAACCCAGAGAGATCC 1800
QY 1815 ATTTAAAGAGAAACTTCTAGAAATTAATGACAAGGATTCAAACTTATTGTCAAATGAGTCC 1874
Db 1801 ATTTAAAGAGAAACTTCTAGAAATTAATGACAAGGATTCAAACTTATTGTCAAATGAGTCC 1860
QY 1875 AATGTCAGATTTTGGAACTCAAACCTATGAACAAATGGGCCATTCAAATGGAAAAAAAAGC 1934
Db 1861 AATGTCAGATTTTGGAACTCAAACCTATGAACAAATGGGCCATTCAAATGGAAAAAAAAGC 1920
QY 1935 TGCAAAAAAAGGAAATCGCAAGAAACGTTGTTGTGCAAGAACATTTGAGGAAGTACAATGA 1994
Db 1921 TGCAAAAAAAGGAAATCGCAAGAAAGTGTGTTGTGCAAGAACATTTGAGGAAGTACAATAA 1980
QY 1995 GGCCCTACAAATTAATGACACAAATTCGAATGATAGATGCGTATACTCATCTTGAAAACTTT 2054
Db 1981 GGCCCTACAAATTAATGACACAAATTCGAATGATAGATGCGTATACTCATCTTGAAAACTTT 2040
QY 2055 CTATAATGAAGAGAAAGATAAGAAAGTTTGCACTCATAGAAGATGATAGTATGAGGGTGG 2114
Db 2041 CTATAATGAAGAGAAAGATAAGAAAGTTTGCACTCATAGAAGATGATAGTATGAGGGTGG 2100
QY 2115 TGATGATGAGTATTGTGATGGTGTGATGAAGATGAGGATGATTTAAAGAAAACTTTGAAAACT 2174
Db 2101 TGATGATGAGTATTGTGATGGTGTGATGAAGATGAGGATGATTTAAAGAAAACTTTGAAAACT 2160
QY 2175 GGATGAAAACAGATAGATTCTCATGACTTTTATTTTGGAAAAACAATAAAATGTTGAAAAAG 2234
Db 2161 GGATGAAAACAGATAGATTCTCATGACTTTTATTTTGGAAAAACAATAAAATGTTGAAAAAG 2220
QY 2235 GCTGGCTGAAAAACCCAGAAATATGAAAATGAAAAGCTGACCAAAATTAAGAAAAATACCATAAT 2294
Db 2221 GCTGGCTGAAAAACCCAGAAATATGAAAATGAAAAGCTGACCAAAATTAAGAAAAATACCATAAT 2280
QY 2295 GGAGCAATATACCTAGGACTGAGGAATCAGCACGAGGAATAATCTTTTACAAAAACACGACA 2354
Db 2281 GGAGCAATATACCTAGGACTGAGGAATCAGCACGAGGAATAATCTTTTACAAAAACACGACA 2340
QY 2355 GAGTGCATATCGCTTTCCAGTGGATTACTGAAAAATGAAAAATTTGCTGAGAGTAGGAGT 2414
Db 2341 GAGTGCATATCGCTTTCCAGTGGATTACTGAAAAATGAAAAATTTGCTGAGAGTAGGAGT 2400
QY 2415 CAAAGCCCAACCATCTGATTGGAGCTGGACAAGAGTGAATTCAAACCCCATGACACAGAA 2474
Db 2401 CAAAGCCCAACCATCTGATTGGAGCTGGACAAGAGTGAATTCAAACCCCATGACACAGAA 2460
QY 2475 TGAACAAAAAAGAAAGTCATTAGTAAAAATTTGCACTGGAAAAATCAATCTGCTTATCGCTAC 2534
Db 2461 TGAACAAAAAAGAAAGTCATTAGTAAAAATTTGCACTGGAAAAATCAATCTGCTTATCGCTAC 2520
QY 2535 CACAGTGGCAGAAAGAGTCTTGATATTAAAGAAATGTPAACATTTGTTATCCGTTATGGTCT 2594

Db 2521 CACAGTGGCAGAAGGCTCTGGATATTAAAGAAATGTAACATTGTTATCCGTTATGCTCT 2580

QY 2595 CGTCACCAATGAATAGCCATGGTCCAGGCCCGTGGTCGAGCCAGAGCTGATGAGAGCAC 2654

Db 2581 CGTCACCAATGAATAGCCATGGTCCAGGCCCGTGGTCGAGCCAGAGCTGATGAGAGCAC 2640

QY 2655 CTACGTCCTGTTGCTCACAGTGGTTCAGGAGTTATCGAACATGAGACAGTTAATGATTT 2714

Db 2641 CTACGTCCTGTTGCTCACAGTGGTTCAGGAGTTATCGAACGTCGAGACAGTTAATGATTT 2700

QY 2715 CCGAGAGAAGATGATGTATAAAGCTATACATTTGTTCAAAATATGAAACCAGAGGAGTA 2774

Db 2701 CCGAGAGAAGATGATGTATAAAGCTATACATTTGTTCAAAATATGAAACCAGAGGAGTA 2760

QY 2775 TGCTCATAAGATTTTGGAAATTACAGATGCAAAAGTATAATGGAAGAAAAAATGAAACCAA 2834

Db 2761 TGCTCATAAGATTTTGGAAATTACAGATGCAAAAGTATAATGGAAGAAAAAATGAAACCAA 2820

QY 2835 GAGAAATATTGCCAAGCATTACAAGATAAACCCATCACATAACTTTCCTTTGGCAAAA 2894

Db 2821 GAGAAATATTGCCAAGCATTACAAGATAAACCCATCACATAACTTTCCTTTGGCAAAA 2880

QY 2895 CTGCAGTGTCTAGCCTGTTCTGGGGAAGATATCCATGTAATGAGAAAAATGCATCACGT 2954

Db 2881 CTGCAGTGTCTAGCCTGTTCTGGGGAAGATATCCATGTAATGAGAAAAATGCATCACGT 2940

QY 2955 CAATATGACCCAGAAATCAAGGAACCTTACATTTGTAAGAGAAAAAACAAGCACCTGCAAAA 3014

Db 2941 CAATATGACCCAGAAATCAAGGAACCTTACATTTGTAAGAGAAAAAACAAGCACCTGCAAAA 3000

QY 3015 GAAGTGTCCGACTATCAATAAATGGTGAAATCATCTGCAAAATGTGGCCAGGCTTGGGG 3074

Db 3001 GAAGTGTCCGACTATCAATAAATGGTGAAATCATCTGCAAAATGTGGCCAGGCTTGGGG 3060

QY 3075 AACAATGATGGTGCACAAAGGCTTAGATTGCTTGCCTTGTCTCAAATAAGGAATTTTGTAGT 3134

Db 3061 AACAATGATGGTGCACAAAGGCTTAGATTGCTTGCCTTGTCTCAAATAAGGAATTTTGTAGT 3120

QY 3135 GGTTTTCAAAAATAATTCAACAAAGAAACAATAACAAAAAGTGGGTAGAATTAACCTATCAC 3194

Db 3121 GGTTTTCAAAAATAATTCAACAAAGAAACAATAACAAAAAGTGGGTAGAATTAACCTATCAC 3180

QY 3195 ATTTCCCAATCTTGACTATTTCAGAAATGCTGTTTATTTAGTGATGAGGATTAGCACTTGAT 3254

Db 3181 ATTTCCCAATCTTGACTATTTCAGAAATGCTGTTTATTTAGTGATGAGGATTAGCACTTGAT 3240

QY 3255 TGAAGATCTTTTAAAAATACTATCAGTTAAACATTTTAATATGATTATGATTAAATGATTC 3314

Db 3241 TGAAGATCTTTTAAAAATACTATCAGTTAAACATTTTAATATGATTATGATTAAATGATTC 3300

QY 3315 ATTATGCTACAGAACTGACATAAGATCAATAAAATGATTGTTTACTCT 3364

Db 3301 ATTATGCTACAGAACTGACATAAGATCAATAAAATGATTGTTTACTCT 3350

RESULT 4

AY017378

LOCUS

DEFINITION Homo sapiens RNA helicase-DEAD box protein RH116 mRNA, linear PRI 28-NOV-2003

cds.

AY017378

VERSION AY017378.1 GI:12621065

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Cocude, C., Truong, M.-J., Billaut-Mulot, O., Delsart, V., Darcissac, E., Capron, A., Mouton, Y. and Bahr, G.M.

TITLE A novel cellular RNA helicase, RH116, differentially regulates cell growth, programmed cell death and human immunodeficiency virus type 1 replication

J. Gen. Virol. 84 (12), 3215-3225 (2003)

2 (bases 1 to 3373)

Cocude, C., Kolesnitchenko, V., Billaut-Mulot, O., Truong, M.-J., Capron, A. and Bahr, G.M.

Identification of a new RNA helicase (RH116) regulated by the immunomodulator Murabutide

Unpublished

3 (bases 1 to 3373)

Cocude, C., Kolesnitchenko, V., Billaut-Mulot, O., Truong, M.-J., Capron, A. and Bahr, G.M.

Direct Submission

Submitted (09-JAN-2001) Laboratoire d'Immunologie Moleculaire de l'Infection et de l'Inflammation, Institut Pasteur de Lille, 1 rue du Professeur Calmette BP 245, Lille 59019, France

Location/Qualifiers

1. .3373

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

155..3232

/codon_start=1

/product="RNA helicase-DEAD box protein RH116"

/protein_id="AAG54076.1"

/db_xref="GI:12621066"

/translation="MSGYSTDENFRYLISCFRVRVMYIQVEPVLVDYLPFLPAEVKEIQRTVATSGNQAVELLSTLEKGVHLGWTRFVEALRRRTGTSPLAARYMNPDLTDLPSFENAHDEYQLNLLQPTLVLDKLLVRDVKMCEEELLTIEDRNRIAAENNGNESGVRELLKRIVQENWFSALNLVLRQTGNNELVQELTSDCSNAEINLSQVDGPQVEBQLSTTVQPNLEKVGWGMENSSSESPADSSVSESDTSLAEGSVSCLDESIGHNSNMGSDSGTMGSDSDENVAARASPEPELQRPYQMEVAQPALEGNIIICLPTGSGKTRVAVYIAKDHLDKKKASEPKVIVLVNKKLVLEQLFRKEFPFLKKWYRVIGLSGDTQLKISFPEVVKSDIIISTAQILENSLNLNLENGEDAGVOLSDFSFIIDECHHTNK EAVYNNIMRHYLMOKLKNRLKKENKVPVLPQILGLTASPGVGATQAKAAEEHILKLCANLDAFTIKTVKENLDQKNQIQEPCKFAIADATREDPFKEKLEIMTRIQTVCQMSPMSDFGTQPYEQWAIQMEKAAKEGNRKESVCAEHLRKYNAKLQINDTIRMIDAYTHLETFFYNEEKDKKFAVIEDSDDEGGDDDEYCDGDEDEDLKKPLKLDETDRFLMTLFFE NNKMLKRLAENPEYENEKLTCLRNTIMEQYTRTEESARGIIFTKTRQSAVALSQWITENEKFAEVGVKAHLIGAGHSSEFKPMTQNEQKEVISKFKTGKINLLIATTVABEGLDIKECNIVIRYGLVTNEIAMVQARGRARADESTYVLVAHSGSVIERETVNDPFRKMMYKAIHCVQNMKPEEYAHKILELQMSIMEKMKTKRNI AKHYKNPNSLITFLCKNCVSLA CSGEDIHVIEKMHVNMTPFEKELYIVRENKALQKCADYQINGEIIICKGQAWGTMMVHKGDLPLCLKIRNFVVVFKNSTKKQYKKWVELPITFPNLDYSECCLFSDDED"

1148..1201

/note="Region: DEAD-like helicase domain"

1481..1627

/note="Region: DEAD/DEAH box"

2606..2647

/note="conserved C-terminal helicase region; Region: helicase C"

ORIGIN

Query Match 99.2%; Score 3338.8; DB 9; Length 3373;

Best local Similarity 99.8%; Pred. No. 0;

Matches 3343; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 15 GAGCCCTGTGGACAAACCTCGTCATTGTTCAGGCACAGAGCGGTAGACCCCTGCTTCTTAAG 74

Db 1 GGGCCCTGTGGACAAACCTCGTCATTGTTCAGGCACAGAGCGGTAGACCCCTGCTTCTTAAG 60

QY 75 TGGGCAGCGGACAGCGGCACGACATTTTACCTGTCCGCAGACAACAGCACCATCTGCT 134

Db 61 TGGGCAGCGGACAGCGGCACGACATTTTACCTGTCCGCAGACAACAGCACCATCTGCT 120

QY 135 TGGGAGAACCCCTCTCCCTTCTCTGAGAAAGAAAGATGTGGAATGGGTATTCACAGACGA 194

Db 121 TGGGAGAACCCCTCTCCCTTCTCTGAGAAAGAAAGATGTGGAATGGGTATTCACAGACGA 180

QY 195 GAATTTCCGCTATCTCATCTCGTGCTTCAGGGCCAGGGTGAATAATGTACATCCAGGTGGA 254

Db 181 GAATTTCCGCTATCTCATCTCGTGCTTCAGGGCCAGGGTGAATAATGTACATCCAGGTGGA 240

QY 255 GCCTGTGTGGACTACCTGACCTTTCTGCCTGCAGAGGTGAAGGACGAGATTTCAGAGGAC 314

Db 241 GCCTGTCTGGACTACCTGACCTTTCTGCCTGCAGAGGTGAAGGAGCAGATTTCAGAGGAC 300

Qy 315 AGTGCACCTCCGGGAACATGCAGGCAGTTGAACTGCTGCTGAGCACTTTGGAGAAGGG 374

Db 301 AGTCGCACCTCCGGGAACATGCAGGCAGTTGAACTGCTGCTGAGCACTTTGGAGAAGGG 360

Qy 375 AGTCTGGCACCTTGGTTGGACTCGGGGAATTCTGTGAGGCCCTCCGGAGAACCGGCAGCCC 434

Db 361 AGTCTGGCACCTTGGTTGGACTCGGGGAATTCTGTGAGGCCCTCCGGAGAACCGGCAGCCC 420

Qy 435 TCTGGCCGCCGCTACATGAACCTGAGCTCACGGACTTGCCTCTCCATCGTTTGAGAA 494

Db 421 TCTGGCCGCCGCTACATGAACCTGAGCTCACGGACTTGCCTCTCCATCGTTTGAGAA 480

Qy 495 CGCTCATGATGAATATCTCCAACCTGCTGAACCTCCTTCAGCCCCACTCTGGTGGACAAGCT 554

Db 481 CGCTCATGATGAATATCTCCAACCTGCTGAACCTCCTTCAGCCCCACTCTGGTGGACAAGCT 540

Qy 555 TCTAGTTAGAGACGTCTTGGATAAGTGCATGGAGGAGAACTGTTGACAAATTGAAGACAG 614

Db 541 TCTAGTTAGAGACGTCTTGGATAAGTGCATGGAGGAGAACTGTTGACAAATTGAAGACAG 600

Qy 615 AAACCGGATTGCTGCTGCAGAAAACAATGGAATGAATCAGGTGTAAGAGAGCTACTAAA 674

Db 601 AAACCGGATTGCTGCTGCAGAAAACAATGGAATGAATCAGGTGTAAGAGAGCTACTAAA 660

Qy 675 AAGGATTGTCAGAAAAGAAACTGGTCTCTGCAATTTCTGAATGTTCTTCGTCAAAACAGG 734

Db 661 AAGGATTGTCAGAAAAGAAACTGGTCTCTGCAATTTCTGAATGTTCTTCGTCAAAACAGG 720

Qy 735 AAACAAATGAACCTTGTCCAAGAGTTAACAGGCTCTGATTGCTCAGAAAAGCAATGCAGAGAT 794

Db 721 AAACAAATGAACCTTGTCCAAGAGTTAACAGGCTCTGATTGCTCAGAAAAGCAATGCAGAGAT 780

Qy 795 TGAGAAATTTATCAAGATTGATGGTCTCAAGTGAAGAGCAACTTCTTTCAACACACAGT 854

Db 781 TGAGAAATTTATCAAGATTGATGGTCTCAAGTGAAGAGCAACTTCTTTCAACACACAGT 840

Qy 855 TCAGCCAAATCTGGAGAAGGAGGTCTGGGCGATGGAGAATAACTCATCAGAATCATCTTT 914

Db 841 TCAGCCAAATCTGGAGAAGGAGGTCTGGGCGATGGAGAATAACTCATCAGAATCATCTTT 900

Qy 915 TGCAGATTCTTCTGTAGTTTCAGAAATCAGACACAAGTTTGGCAGAGGAAGTGTCTAGCTG 974

Db 901 TGCAGATTCTTCTGTAGTTTCAGAAATCAGACACAAGTTTGGCAGAGGAAGTGTCTAGCTG 960

Qy 975 CTTAGATGAAAGTCTTGACATAACAGCAACATGGGCAGTGTATTCAGGCACCATGGGAAG 1034

Db 961 CTTAGATGAAAGTCTTGACATAACAGCAACATGGGCAGTGTATTCAGGCACCATGGGAAG 1020

Qy 1035 TGATTCAGATGAAGAGAAATGTGGCAGCAAGAGATCCCCGGAGCCAGAACTCCAGCTCAG 1094

Db 1021 TGATTCAGATGAAGAGAAATGTGGCAGCAAGAGATCCCCGGAGCCAGAACTCCAGCTCAG 1080

Qy 1095 GCCTTACCAAATGGAAGTTGCCAGCCAGCCCTTGGAGGGGAAGAAATATCATCATCTGCCT 1154

Db 1081 GCCTTACCAAATGGAAGTTGCCAGCCAGCCCTTGGAGGGGAAGAAATATCATCATCTGCCT 1140

Qy 1155 CCCTACAGGGAGTGGAAAAACCAGAGTGGCTGTTTACATTCGCCAGGATCACTTAGACAA 1214

Db 1141 CCCTACAGGGAGTGGAAAAACCAGAGTGGCTGTTTACATTCGCCAGGATCACTTAGACAA 1200

Qy 1215 GAAGAAAAAAGCATCTGAGCCTGGAAAAAGTTATAGTTCTTGTCAATAAGGTACTGCTAGT 1274

Db 1201 GAAGAAAAAAGCATCTGAGCCTGGAAAAAGTTATAGTTCTTGTCAATAAGGTACTGCTAGT 1260

Qy 1275 TGAACAGCTCTCCGCAAGGAGTTCCAACCATTTTGAAGAAATGGTATCGTGTATTGG 1334

Db 1261 TGAACAGCTCTCCGCAAGGAGTTCCAACCATTTTGAAGAAATGGTATCGTGTATTGG 1320

Qy 1335 ATTAAGTGTGATACCCAACTGAAAATATCATTTCCAGAAGTTGTCAAGTCTGTGATAT 1394

Db 1321 ATTAAGTGTGATACCCAACTGAAAATATCATTTCCAGAAGTTGTCAAGTCTGTGATAT 1380

Qy 1395 TATTATCAGTACAGCTCAAATCCTTTGAAAACTCCCTCTTAAACTTTGAAAAATGGAGAAGA 1454

Db 1381 TATTATCAGTACAGCTCAAATCCTTTGAAAACTCCCTCTTAAACTTTGAAAAATGGAGAAGA 1440

Qy 1455 TGCTGGTGTTCAAATTGTCTAGACTTTTCCCTCATTTATCATTTGATGATGTCTACACCAA 1514

Db 1441 TGCTGGTGTTCAAATTGTCTAGACTTTTCCCTCATTTATCATTTGATGATGTCTACACCAA 1500

Qy 1515 CAAAGAAGCAGTGTATAATAACATCATGAGGCATTTATTTGATGCAGAAAGTTGAAAAACAA 1574

Db 1501 CAAAGAAGCAGTGTATAATAACATCATGAGGCATTTATTTGATGCAGAAAGTTGAAAAACAA 1560

Qy 1575 TAGACTCAAGAAAGAAAAACAACCAAGTGTATTCCTTCCCTCAGATACTGGGACTAACAGC 1634

Db 1561 TAGACTCAAGAAAGAAAAACAACCAAGTGTATTCCTTCCCTCAGATACTGGGACTAACAGC 1620

Qy 1635 TTCACCTGGTGTGGAGGGGCCAGCAAGCAAGCCAAAGCTGAAGAACAACATTTTAAACT 1694

Db 1621 TTCACCTGGTGTGGAGGGGCCAGCAAGCAAGCTGAAGAACAACATTTTAAACT 1680

Qy 1695 ATGTGCCAATCTTGTGATGCAATTTACTATTAAACTGTTTAAAGAAAACTTTGATCAACTGAA 1754

Db 1681 ATGTGCCAATCTTGTGATGCAATTTACTATTAAACTGTTTAAAGAAAACTTTGATCAACTGAA 1740

Qy 1755 AAACCAAATACAGGAGCCATGCAAGAAAGTTTGCCATTGCAGATGCAACCAAGAGAGATCC 1814

Db 1741 AAACCAAATACAGGAGCCATGCAAGAAAGTTTGCCATTGCAGATGCAACCAAGAGAGATCC 1800

Qy 1815 ATTTAAAGAGAAACTTCTAGAAATAATGAACAAGGATTCAAACTTATTTGTCAAATGAGTCC 1874

Db 1801 ATTTAAAGAGAAACTTCTAGAAATAATGAACAAGGATTCAAACTTATTTGTCAAATGAGTCC 1860

Qy 1875 AATGTCAGATTTTGGAACTCAACCTCATGAACAATGGGCCATTCAAATGGAAAAAAAGC 1934

Db 1861 AATGTCAGATTTTGGAACTCAACCTCATGAACAATGGGCCATTCAAATGGAAAAAAAGC 1920

Qy 1935 TGCAAAAAAGGAATCGCAAGAAACCTGTTTGTGCAGAACATTTGAGGAAGTACAATGA 1994

Db 1921 TGCAAAAAAGGAATCGCAAGAAACCTGTTTGTGCAGAACATTTGAGGAAGTACAATGA 1980

Qy 1995 GGCCCTACAAATTAATGACACAATTCGAATGATAGATGCGTATPACTCATCTTGAAACTTT 2054

Db 1981 GGCCCTACAAATTAATGACACAATTCGAATGATAGATGCGTATPACTCATCTTGAAACTTT 2040

Qy 2055 CTATAATGAAGAGAAAGATAAGAAAGTTTGCAGTCATAGAAGATGATGATGAGGGTGG 2114

Db 2041 CTATAATGAAGAGAAAGATAAGAAAGTTTGCAGTCATAGAAGATGATGATGAGGGTGG 2100

Qy 2115 TGATGATGAGTATTGTGATGGTGTGATGAAGATGAGGATGATTTAAAGAAAACTTTGAAACT 2174

Db 2101 TGATGATGAGTATTGTGATGGTGTGATGAAGATGAGGATGATTTAAAGAAAACTTTGAAACT 2160

Qy 2175 GGATGAAAACAGATAGATTTCTCATGACTTTTATTTTGGAAAAACAATAAAAAATTTGAAAAAG 2234

Db 2161 GGATGAAAACAGATAGATTTCTCATGACTTTTATTTTGGAAAAACAATAAAAAATTTGAAAAAG 2220

Qy 2235 GCTGGCTGAAAAACCCAGAAATATGAAAAATGAAAAGCTGACCAAAATTAAGAAATACCATAAT 2294

Db 2221 GCTGGCTGAAAAACCCAGAAATATGAAAAATGAAAAGCTGACCAAAATTAAGAAATACCATAAT 2280

Qy 2295 GGAGCAATATACTAGGACTGAGGAATCAGCACGAGGAATAATCTTTACAAAAACACGACA 2354

Db 2281 GGAGCAATATACTAGGACTGAGGAATCAGCACGAGGAATAATCTTTACAAAAACACGACA 2340

Qy 2355 GAGTGCATATCGCCTTTCCAGTGGATTACTGAAAAATGAAAAATTTGCTGTAAGTAGGAGT 2414

Db 2341 GAGTGCATATCGCCTTTCCAGTGGATTACTGAAAAATGAAAAATTTGCTGTAAGTAGGAGT 2400

Qy 2415 CAAAGCCCAACCATCTGATTGGAGCTGGACACAGCAGTGAAGTTCAAACCCCATGACACAGAA 2474

Db 2401 CAAAGCCCAACCATCTGATTGGAGCTGGACACAGCAGTGAAGTTCAAACCCCATGACACAGAA 2460

QY	2475	TGAACAAAAAGAGTCATTAGTAAATTTCCGACTGGAAAAATCAATCTGCTTATCGCTAC	2534
Db	2461	TGAACAAAAAGAGTCATTAGTAAATTTCCGACTGGAAAAATAAATCTGCTTATCGCTAC	2520
QY	2535	CACAGTGGCAGAGAAGGCTCTGGATATTAAAGAAATGTAACATTGTTATCCGTTATGGTCT	2594
Db	2521	CACAGTGGCAGAGAAGGCTCTGGATATTAAAGAAATGTAACATTGTTATCCGTTATGGTCT	2580
QY	2595	CGTCACCAATGAATAGCCATGCTCCAGGCCCGTGGTCGAGCCAGAGCTGATGAGAGCAC	2654
Db	2581	CGTCACCAATGAATAGCCATGCTCCAGGCCCGTGGTCGAGCCAGAGCTGATGAGAGCAC	2640
QY	2655	CTACGTCCTGGTTGCTCACAGTGGTTCAGGAGTTATCGAACATGAGACAGTTAATGATTT	2714
Db	2641	CTACGTCCTGGTTGCTCACAGTGGTTCAGGAGTTATCGAACGTCGAGACAGTTAATGATTT	2700
QY	2715	CCGAGAGAAGATGATGTATAAAGCTATACATTGTGTTCAAATATGAAACCAGAGGAGTA	2774
Db	2701	CCGAGAGAAGATGATGTATAAAGCTATACATTGTGTTCAAATATGAAACCAGAGGAGTA	2760
QY	2775	TGCTCATAGATTTTGGAAATTACAGATGCAAGTATAATGGAAGAAAAATGAAAAACCA	2834
Db	2761	TGCTCATAGATTTTGGAAATTACAGATGCAAGTATAATGGAAGAAAAATGAAAAACCA	2820
QY	2835	GAGAAATATTGCCAAGCATTACAAGATAACCCATCACTAATAACTTTCCTTTGCAAAAA	2894
Db	2821	GAGAAATATTGCCAAGCATTACAAGATAACCCATCACTAATAACTTTCCTTTGCAAAAA	2880
QY	2895	CTGCAGTGTGCTAGCTGTTCTGGGAAGATATCCATGTAATTGAGAAAAATGCATCAGT	2954
Db	2881	CTGCAGTGTGCTAGCTGTTCTGGGAAGATATCCATGTAATTGAGAAAAATGCATCAGT	2940
QY	2955	CAATATGACCCCGAATTCAGGAACTTTACATTGTAAGAGAAAAACAAGCAGTCGAAAA	3014
Db	2941	CAATATGACCCCGAATTCAGGAACTTTACATTGTAAGAGAAAAACAAGCAGTCGAAAA	3000
QY	3015	GAAGTGTCCGACTATCAATAAATGGTGAAATCATCTGCAATGTGGCCAGGCTTGGGG	3074
Db	3001	GAAGTGTCCGACTATCAATAAATGGTGAAATCATCTGCAATGTGGCCAGGCTTGGGG	3060
QY	3075	AACAAATGATGGTGCAAAAGGCTTAGATTTGCCCTTGTCTCAAATAAGGAATTTTGTAGT	3134
Db	3061	AACAAATGATGGTGCAAAAGGCTTAGATTTGCCCTTGTCTCAAATAAGGAATTTTGTAGT	3120
QY	3135	GGTTTTCAAAAAATTAATCAACAAAGAAACAATACAAAAAGTGGGTAGAATTACCTATCAC	3194
Db	3121	GGTTTTCAAAAAATTAATCAACAAAGAAACAATACAAAAAGTGGGTAGAATTACCTATCAC	3180
QY	3195	ATTTCCCAATCTTGACTATTTCAGAAATGCTGTTTATTATTAGTGATGAGGATTAGCATTGAT	3254
Db	3181	ATTTCCCAATCTTGACTATTTCAGAAATGCTGTTTATTATTAGTGATGAGGATTAGCATTGAT	3240
QY	3255	TGAAGATTCTTTTAAAAATACTATCAGTTAAACATTTAATATGATTATGATTATGATTTC	3314
Db	3241	TGAAGATTCTTTTAAAAATACTATCAGTTAAACATTTAATATGATTATGATTATGATTTC	3300
QY	3315	ATTATGCTACAGAACTGACATAAGAATCAATAAAATGATTGTTTACTCT	3364
Db	3301	ATTATGCTACAGAACTGACATAAGAATCAATAAAATGATTGTTTACTCT	3350

RESULT 5
BC078180
LOCUS BC078180 2540 bp mRNA linear PRI 27-JUL-2004
DEFINITION Homo sapiens interferon induced with helicase C domain 1, mRNA (cdna clone IMAGE:6502757), partial cds.
ACCESSION BC078180
VERSION BC078180.1 GI:50415794
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS

1 (bases 1 to 2540)
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shermen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettaman,M., Madan,A., Rodriguez,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinaki,M.I., Skalska,U., Smailus,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002) 12477932

TITLE

JOURNAL
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

2 (bases 1 to 2540)
Strausberg,R.
Direct Submission
Submitted (19-JUL-2004) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@nhgri.nih.gov

REMARK
COMMENT

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 174 Row: p Column: 16
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 27886567.
Location/Qualifiers
1. .2540
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6502757"
/tissue_type="Uterus, leiomyosarcoma"
/clone_lib="NIH_MGC_71"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1. .>2540
/gene="IFIH1"
/note="synonyms: MDA5, MDA-5, Hlcd"
/db_xref="LocusID:64135"
/db_xref="MIM:606951"
224. .>2540
/gene="IFIH1"
/codon_start=1

FEATURES
source

1. .2540
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6502757"
/tissue_type="Uterus, leiomyosarcoma"
/clone_lib="NIH_MGC_71"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1. .>2540
/gene="IFIH1"
/note="synonyms: MDA5, MDA-5, Hlcd"
/db_xref="LocusID:64135"
/db_xref="MIM:606951"
224. .>2540
/gene="IFIH1"
/codon_start=1

CDS

/product="IFIH1 protein"
/protein_id="AAH78180.1"
/db_xref="GI:50415795"
/db_xref="LocusID:64135"
/db_xref="MIM:606951"
translation="MSGYSTDENFRYLISCFRARVRKMYIQVEPVLDTFLPAEVEK
QIQRVATSGNMQAVELLSTLEKGVHLGWTRFEVLEALRRTPGSLAARYMNPETLDL
PSPSPENAHDEYQLNLQPTLVLDKLVLDVLDKCMEEELLTIEDNRRIAAAEENNGN
ESGVRELLKRIQVKNWFSAPLNLVRLQTNNELVQELTSGDCSESNABEINLSQVDDGP
QVEEQLLSTTVQPNLEKEVMGNNSSSESPADSVSVSEDTSLAEGSVSCLDESGLH
NSNMGSDSGTMGSDSEENVAAARSPPELQRLPYQMEVAQPALEGNIIICLPTGSG
KTRVAVYIAKDHLDDKKKASEPGKVIVLNVKLLVEQLFRKEPQPFLLKWRVILGSG
DTQLKISFPEVVKSCDIIISTAQILENSLNLENGEDAGVQLSDFSLLIIDECHTNK
EAVYNNIMRHYLMQKLNRLKKNRPVILPQILGLTASPGVGGATQKAKAEHILK
LCANLDAFTIKTVKENLDQLKNQIQEPCPKFAIADATREDPFKEKLLEIMTRIQYCO
MSPMSDFGTPYEQWAIQMEKKAKEGNRKERVCAEHLRKVNEALQINDTIRMIDAYT
HLETIFYNEEKDKKFAVIEDSDDEGDDDEYCDGDEDDLLKPLKLDDETRFLMTLFE
NNKMLKRLAENPEYENEKLTKLNTIMEQYTRTEESARGIIFTKTRQSAVALSQWITE
NEKFAEVGVKAHHLIGAGHSSEFKPMTQKKK"

ORIGIN

Query Match		73.6%;	Score 2476;	DB 9;	Length 2540;
Best Local Similarity		99.8%;	Pred. No. 0;		
Matches 2479;		Conservative	0;	Mismatches	5;
				Indels	0;
				Gaps	0;
QY	1	GCGCGCCGGCCTGAGAGCCCTGTGGACAACTCGTCATTTGTCAGGCACAGAGCGGTAGAC	60		
Db	56	GCGCGCCGGCCTGAGAGCCCTGTGGACAACTCGTCATTTGTCAGGCACAGAGCGGTAGAC	115		
QY	61	CCTGCTTCTTAAGTGGGAGAACCTCTCCCTTCTCTGAGAAAGAAAGATGTCGAATGGG	120		
Db	116	CCTGCTTCTTAAGTGGGAGAACCTCTCCCTTCTCTGAGAAAGAAAGATGTCGAATGGG	175		
QY	121	CAGCACCATCTGCTTGGGAGAACCTCTCCCTTCTCTGAGAAAGAAAGATGTCGAATGGG	180		
Db	176	CAGCACCATCTGCTTGGGAGAACCTCTCCCTTCTCTGAGAAAGAAAGATGTCGAATGGG	235		
QY	181	TATTCACAGACGAGAAATTCGGCTATCTCATCTCGTCTGCTTCCAGGGCCAGGGTGAATG	240		
Db	236	TATTCACAGACGAGAAATTCGGCTATCTCATCTCGTCTGCTTCCAGGGCCAGGGTGAATG	295		
QY	241	TACATCCAGGTGGAGCCTGTGCTGGACTACCTGACCTTCTGCTGTCAGAGGAGGAG	300		
Db	296	TACATCCAGGTGGAGCCTGTGCTGGACTACCTGACCTTCTGCTGTCAGAGGAGGAG	355		
QY	301	CAGATTGAGGACAGTCCGACCTCCGGGAAACATGAGGAGTGAACCTGCTGCTGAGC	360		
Db	356	CAGATTGAGGACAGTCCGACCTCCGGGAAACATGAGGAGTGAACCTGCTGCTGAGC	415		
QY	361	ACCTTGGAGAAGGAGTCTGGACCTTGGTTGGACTCGGGAATTCCTGGAGGCCCTCCGG	420		
Db	416	ACCTTGGAGAAGGAGTCTGGACCTTGGTTGGACTCGGGAATTCCTGGAGGCCCTCCGG	475		
QY	421	AGAACCGGACGCCCTCTGGCCGCCCGCTACATGAACCTGAGCTCAGGACTTGCCCTCT	480		
Db	476	AGAACCGGACGCCCTCTGGCCGCCCGCTACATGAACCTGAGCTCAGGACTTGCCCTCT	535		
QY	481	CCATCGTTTGAAGAACGCTCATGATGAATATCTCCAACTGCTGAACCTCTCAGCCCCACT	540		
Db	536	CCATCGTTTGAAGAACGCTCATGATGAATATCTCCAACTGCTGAACCTCTCAGCCCCACT	595		
QY	541	CTGGTGGACAAGCTTCTAGTTAGAGACGTCCTTGGATAAGTGCATGGAGGAGAACTGTTG	600		
Db	596	CTGGTGGACAAGCTTCTAGTTAGAGACGTCCTTGGATAAGTGCATGGAGGAGAACTGTTG	655		
QY	601	ACAAATTGAAGACAGAAAACCGGATTGCTGTCAGAAAACAAATGGAAATGAATCAGGTGTA	660		
Db	656	ACAAATTGAAGACAGAAAACCGGATTGCTGTCAGAAAACAAATGGAAATGAATCAGGTGTA	715		
QY	661	AGAGAGCTACTTAAAAAGGATTGTGAGAAAAGAAAACCTGGTTCTCTGCAATTTCTGAATGTT	720		
Db	716	AGAGAGCTACTTAAAAAGGATTGTGAGAAAAGAAAACCTGGTTCTCTGCAATTTCTGAATGTT	775		

QY	721	CTTCGTCAAACAGGAAACAATGAACCTTGTCTCAAGAGTTAACAGGCTCTGATTTGCTCAGAA	780
Db	776	CTTCGTCAAACAGGAAACAATGAACCTTGTCTCAAGAGTTAACAGGCTCTGATTTGCTCAGAA	835
QY	781	AGCAATGCAGAGATTGAGAAATTTATCACAAGTTGATGGTCTCTCAAGTGAAGAGCAACTT	840
Db	836	AGCAATGCAGAGATTGAGAAATTTATCACAAGTTGATGGTCTCTCAAGTGAAGAGCAACTT	895
QY	841	CTTTCAACCAACAGTTTCAGCCAAATCTGGAGAAGGAGGTCTGGGGCATGGAGATAACTCA	900
Db	896	CTTTCAACCAACAGTTTCAGCCAAATCTGGAGAAGGAGGTCTGGGGCATGGAGATAACTCA	955
QY	901	TCAGAATCATCTTTTGCAGATTCTTCTGTAGTTTCAGAATCAGACACAAGTTTGGCAGAA	960
Db	956	TCAGAATCATCTTTTGCAGATTCTTCTGTAGTTTCAGAATCAGACACAAGTTTGGCAGAA	1015
QY	961	GGAAGTGTGAGCTGCTTAGATGAAAAGTCTTGGACATAAACAGCAACATGGGCAGTATTCA	1020
Db	1016	GGAAGTGTGAGCTGCTTAGATGAAAAGTCTTGGACATAAACAGCAACATGGGCAGTATTCA	1075
QY	1021	GGCACCATGGGAAGTGATTTCAGATGAAGAGAATGTGGCAGCAAGAGCATCCCCGGAGCCA	1080
Db	1076	GGCACCATGGGAAGTGATTTCAGATGAAGAGAATGTGGCAGCAAGAGCATCCCCGGAGCCA	1135
QY	1081	GAACTCCAGCTCAGGCCCTTACCAATGGAAGTTGCCAGCCAGCCTTGGAAAGGAAGAAT	1140
Db	1136	GAACTCCAGCTCAGGCCCTTACCAATGGAAGTTGCCAGCCAGCCTTGGAAAGGAAGAAT	1195
QY	1141	ATCATCATCTGCTCCCTACAGGGAGTGAAAAACCCAGAGTGGCTGTTTACATTGCCAAG	1200
Db	1196	ATCATCATCTGCTCCCTACAGGGAGTGAAAAACCCAGAGTGGCTGTTTACATTGCCAAG	1255
QY	1201	GATCACTTAGACAAAGAAAAAGCATCTGAGCCTGGAAAAAGTTATAGTTCTTTGTCAAT	1260
Db	1256	GATCACTTAGACAAAGAAAAAGCATCTGAGCCTGGAAAAAGTTATAGTTCTTTGTCAAT	1315
QY	1261	AAGGTACTGCTAGTTGAACAGCTCTTCGCAAGGAGTTCCAACCATTTTGAAGAAATGG	1320
Db	1316	AAGGTACTGCTAGTTGAACAGCTCTTCGCAAGGAGTTCCAACCATTTTGAAGAAATGG	1375
QY	1321	TATCGTGTATTGGATTAAGTGGTGATACCCAACTGAAAAATATCATTTCCAGAAAGTTGTC	1380
Db	1376	TATCGTGTATTGGATTAAGTGGTGATACCCAACTGAAAAATATCATTTCCAGAAAGTTGTC	1435
QY	1381	AAGTCTGTGATATTATATCAGTACAGTCAAACTCCTTGAAAACTCCCTCTTTAAACTTG	1440
Db	1436	AAGTCTGTGATATTATATCAGTACAGTCAAACTCCTTGAAAACTCCCTCTTTAAACTTG	1495
QY	1441	GAAATGGAGAAGATGCTGGTGTTCAAATGTGAGACTTTTCCCTCATTTATCATTTGATGAA	1500
Db	1496	GAAATGGAGAAGATGCTGGTGTTCAAATGTGAGACTTTTCCCTCATTTATCATTTGATGAA	1555
QY	1501	TGTCATCACACCAACAAAGAGCAGTGTATATAAATCATCATGAGGATTTATTTGATGCAG	1560
Db	1556	TGTCATCACACCAACAAAGAGCAGTGTATATAAATCATCATGAGGATTTATTTGATGCAG	1615
QY	1561	AAGTTGAAAAACAATAGACTCAAGAAAGAAAAACAAACCAAGTATTCCTTCCAGATA	1620
Db	1616	AAGTTGAAAAACAATAGACTCAAGAAAGAAAAACAAACCAAGTATTCCTTCCAGATA	1675
QY	1621	CTGGGACTAACAGCTTCACCTGGTGTGGAGGGGCCACGAAGCAAGCAAGCTGAAGAA	1680
Db	1676	CTGGGACTAACAGCTTCACCTGGTGTGGAGGGGCCACGAAGCAAGCAAGCTGAAGAA	1735
QY	1681	CACATTTTAAAACTATGTGCCAATCTTGTGATGCAATTTACTATTAAAACTGTTAAAAA	1740
Db	1736	CACATTTTAAAACTATGTGCCAATCTTGTGATGCAATTTACTATTAAAACTGTTAAAAA	1795
QY	1741	CTTGATCAACTGAAAAACCAATAACAGGAGCCATGCAAGAGTTTGGCAATTTGCAGATGCA	1800
Db	1796	CTTGATCAACTGAAAAACCAATAACAGGAGCCATGCAAGAGTTTGGCAATTTGCAGATGCA	1855
QY	1801	ACCAGAGAAGATCCATTTTAAAGAGAAACCTTCTAGAAATAATGACAAGGATTTCAAACTTAT	1860


```
|||||
Db 1856 ACCAGAGAAGATCCATTAAAGAGAAACTTCTAGAAATAATGACAAGGATTCAAACTTAT 1915

QY 1861 TGTCAAATGAGTCCAATGTCAGATTTTGGAACTCAACCCATATGAACAATGGGCCATTCAA 1920
|||||
Db 1916 TGTCAAATGAGTCCAATGTCAGATTTTGGAACTCAACCCATATGAACAATGGGCCATTCAA 1975

QY 1921 ATGGAAAAAAGAGCTGCAAAAAAAGGAAATCGCAAAAGAACGTGTTTGTGCAGAACATTTG 1980
|||||
Db 1976 ATGGAAAAAAGAGCTGCAAAAGAGGAAATCGCAAAAGAACGTGTTTGTGCAGAACATTTG 2035

QY 1981 AGGAAGTACAATGAGGCCCTACAAATTAATGACACAATTCGAATGATAGATGCGTATACT 2040
|||||
Db 2036 AGGAAGTACAATGAGGCCCTACAAATTAATGACACAATTCGAATGATAGATGCGTATACT 2095

QY 2041 CATCTTGAACTTTCTATAATGAAGAGAAAGATAAGAAGTTTGCAGTCAATAGAAAGATGAT 2100
|||||
Db 2096 CATCTTGAACTTTCTATAATGAAGAGAAAGATAAGAAGTTTGCAGTCAATAGAAAGATGAT 2155

QY 2101 AGTGATGAGGGTGGTGATGATGAGTATTGTGATGGTGATGAAGATGAGGATGATTTAAAG 2160
|||||
Db 2156 AGTGATGAGGGTGGTGATGATGAGTATTGTGATGGTGATGAAGATGAGGATGATTTAAAG 2215

QY 2161 AAACCTTTGAACTGGATGAAACAGATAGATTTCTCATGACTTTATTTTTTGAAACAAT 2220
|||||
Db 2216 AAACCTTTGAACTGGATGAAACAGATAGATTTCTCATGACTTTATTTTTTGAAACAAT 2275

QY 2221 AAAATGTTGAAAAGCTGGCTGAAAACCCAGAATATGAAAATGAAAAGCTGACCAAAATTA 2280
|||||
Db 2276 AAAATGTTGAAAAGCTGGCTGAAAACCCAGAATATGAAAATGAAAAGCTGACCAAAATTA 2335

QY 2281 AGAAATACCAATAATGGAGCAATATACTAGGACTGAGGAATCAGCACGAGGAATAATCTTT 2340
|||||
Db 2336 AGAAATACCAATAATGGAGCAATATACTAGGACTGAGGAATCAGCACGAGGAATAATCTTT 2395

QY 2341 ACAAAAACACGACAGAGTGATATGCGCTTTCCAGTGGATTACTGAAAATGAAAAATTT 2400
|||||
Db 2396 ACAAAAACACGACAGAGTGATATGCGCTTTCCAGTGGATTACTGAAAATGAAAAATTT 2455

QY 2401 GCTGAAGTAGGAGTCAAAGCCCCACCATCTGATTGGAGCTGGACACAGCAGTGATTCAAA 2460
|||||
Db 2456 GCTGAAGTAGGAGTCAAAGCCCCACCATCTGATTGGAGCTGGACACAGCAGTGATTCAAA 2515

QY 2461 CCCATGACACAGAATGAACAAAA 2484
|||||
Db 2516 CCCATGACACAAAAAANAANAANA 2539
```

```
RESULT 6
CQ777379
LOCUS CQ777379 3771 bp DNA linear PAT 11-MAR-2004
DEFINITION Sequence 1065 from Patent EP1394274.
ACCESSION CQ777379
VERSION CQ777379.1 GI:45380363
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Ohtani,N., Sugita,Y., Yamaya,M., Kubo,H., Nagai,H. and Izuohara,K.
TITLE Methods of testing for bronchial asthma or chronic obstructive
pulmonary disease
JOURNAL Patent: EP 1394274-A 1065 03-MAR-2004;
Genox Research, Inc. (JP)
FEATURES
source
1. .3771
/organism="Mus musculus"
/mol_type="unassigned DNA"
/db_xref="taxon:10090"
ORIGIN
Query Match 65.5%; Score 2202.4; DB 6; Length 3771;
|
```

```
Best Local Similarity 80.7%; Pred. No. 0;
Matches 2609; Conservative 0; Mismatches 616; Indels 7; Gaps 3;

QY 121 CAGCACCATCTGCTTGGGAGAACCTCTCTCCC-TTCTCTGAGAAAGAAAGATGTCGAATGG 179
|||||
Db 186 CCGCCCCCGCCCCCGGGAGAGTCTCTCTCCCATTTCTCTGAGACCTCAACGATGTCGATTGT 245

QY 180 GTATTCCACAGACGAGAAATTTCCGCTATCTCATCTCTGCTTTCAGGGCCAGGGTGAAAAT 239
|||||
Db 246 CTGTTCTGCAGAGGACAGCTTCAGGAATCTCATCTTATTCTTCAGGCCCAGGCTGAAAAT 305

QY 240 GTACATCCAGGTGGAGCCTGTGCTGGACTACCTGACCTTTCTGCCTGCAGAGGTGAAGGA 299
|||||
Db 306 GTACATTCAGGTGGAGCCAGTGTGGACCACCTCATCTTTCTGTCTGCAGAAACCAAAGA 365

QY 300 GCAGATTTCAGAGGACAGTCGCCACCTCCGGGAACATGCAGGCAGTTGAACTGCTGCTGAG 359
|||||
Db 366 GCAGATTCTTAAAGATCAACACCTGTGGTAACACCCAGCGCGGCAGAACTGCTGCTGAG 425

QY 360 CACCTTGGAGAAGGGAGTCTGSCACCTTGGTTGGACTCTGGGAATTCGTGGAGGCCCTCCG 419
|||||
Db 426 CACCTTGGAGCAGGGACAATGGCCTCTGGGATGGACGCAGATGTTCTGTGGAGGCCCTAGA 485

QY 420 GAGAACCGGCAGCCCTCTGGCCGCCCGCTACATGAACCTGAGCTCACGGACTTGCCTC 479
|||||
Db 486 GCACAGTGGCAATCCCTTAGCCGCGGCTATGTCAAACCCACACTCACTGATCTGCCCTC 545

QY 480 TCCATCGTTTGAGAACGCTCATATGATGAATATCTCCAAGTCTGAACTCCTTCAGCCAC 539
|||||
Db 546 TCCCTCCTCTGAGACTGCCCATGACGAGTGTCTCCACTGTCTGACCCCTCCTCCAGCCAC 605

QY 540 TCTGGTGGACAAGCTTCTAGTTAGAGACGTCTTGGATAAGTGCATGGAGGAGGAATGTT 599
|||||
Db 606 TTTGGTGGACAAACTTCTGATTAAAGATGTCTTTGGACACTTGTTCGAGAAGGGACTATT 665

QY 600 GACAATTGAAGACAGAAACCGGATTGCTGCTGCAGAAAAACAATGGAATGAATCAGGTGT 659
|||||
Db 666 GACAGTCGAAGACAGAAATCGGATTTCTGCTGCAGGAAACACGCGGGAATGAGTCAGGTGT 725

QY 660 AAGAGAGCTACTAAAAAGGATTGTGCAGAAAGAAAACGTTCTCTGCATTTCTGCAATGT 719
|||||
Db 726 AAGAGAGCTGCTGAGAAGGATTGTGCAGAAGGAAAACGTTTCTACCTTCTCCTGGATGT 785

QY 720 TCTTCGTCAAAACAGGAAACAATGAACTTGTCCAAGAGTTAAACAGGCTCTGTATTGCTCAGA 779
|||||
Db 786 TCTGCGCCAAAACGTGAAATGATGCACTATTCCAAGAACTAAACAGGTGGAGGCTGCCCAGA 845

QY 780 AAGCAATGCAGAGATTGAGAAATTTATCACAAGTTGATGTTCTCAAGTGGAAAGAGCAACT 839
|||||
Db 846 AGACAACACAGACTTGGCTAACTCGTCTCACAGAGATGGCCTGCAGCTAATGAGTGTCT 905

QY 840 TCTTTCAACCACAGTTTCAGCCCAATCTGGAGAAGGAGGTCTGGGGCATGGAGAATAACTC 899
|||||
Db 906 TCTGCCTGCTGTCGATGAGTCAAGTCTGGAGACAGAGGCTGGAAACGTAGACGACATATT 965

QY 900 ATCAGAAATCATCTTTTGCAGATTCTTCTGTAGTTTTCAGAAATCAGACACAAGTTTGGCAGA 959
|||||
Db 966 ACCAGAGGCTTCTTGTACAGATTCTGTGTGACCACAGAAATCAGACACAAGTTTGGCAGA 1025

QY 960 AGGAAGTGTGAGTGTGTTAGTGAAGTCTTTGGACATAACAGCAACATGGGCGAGTATTC 1019
|||||
Db 1026 AGGAAGTGTGAGTGTGTTAGTGAAGTCTTTGGACATAACAGCAACATGGGCGAGGATTC 1085

QY 1020 AGGCACCATGGGAAGTGAATTCAGATGAA--GAGAAATGTGGCAGCAAGAGCATCCCGGA 1076
|||||
Db 1086 AGGCACCATGGGAAGTGAATTCAGATGAAAGTGTATCCAGACAAAAAAGAGTATCCCCGA 1145

QY 1077 GCCAGAACTCCAGCTCAGGCCCTTACCAAATGGAAGTTGCCCGCAGCCCTTGGAAAGGAA 1136
|||||
Db 1146 GCCAGAACTCCAGCTCAGGCCCTTACCAAATGGAAGTGGCCCAACCCAGCTCTAGATGGGA 1205

QY 1137 GAATATCATCATCTGCCTCCCTACAGGGAGTGGAATAACAGAGTGGCTGTTTACATTGC 1196
|||||
```

Db 1206 GAATATTATTATCTGCCTCCACCGGGAGTGGGAAAACACAGAGTGGCTGTTTACATCAC 1265

QY 1197 CAAGGATCACTTAGACAAGAGAAAAAGCATCTGAGCCTGGAAAAAGTTATAGTTCCTGT 1256

Db 1266 CAAAGATCACTTAGACAAGAGAACGAGGCATCTGAAATCCGGGAAGGTTATCGTTCCTGT 1325

QY 1257 CAATAAGGTACTGCTAGTTGAAACAGCTCTTCCGGAAGGAGTTCCAACCAATTTTGAAGAA 1316

Db 1326 CAATAAGGTAAATGTTAGCAGAAACAACCTTTCCGAAAAGAGTTCAACCCATATTTGAAGAA 1385

QY 1317 ATGGATCGTGTATTGGATTAAAGTGGTGATACCCAACTGAAAAATATCAATTCAGAAAGT 1376

Db 1386 ATGGTATCGAAATTATTGGATTAAAGTGGCGATACCCAGCTGAAAAATATCATTTCCAGAAAGT 1445

QY 1377 TGTCAAAGTCCTGTGATATTATTATCAGTACAGCTCAAACTCCTTGAAAACTCCCTCTTAAA 1436

Db 1446 TGTCAAATCTTACGATGTTATTATCAGCACTGCTCAAACTCCTTGAAAACTCCCTCTTAAA 1505

QY 1437 CTTGGAAAATGGAGAAGATGCTGGTGTTCAAATTGTCAGACTTTTCCCTCATTTATCATTTGA 1496

Db 1506 TCTGGAGAGTGGAGACGATGACGCTGTGCAGCTGTGCAGACTTCTCTCTCATTTATCATTTGA 1565

QY 1497 TGAATGTCATCACACCAACAAAGAACGAGTGTATAATAAATCAATCATGAGGCAATATTGAT 1556

Db 1566 TGAGTGCCATCACACCAACAAAGAGGCAGTCTATAAACACATCATGAGACGATATTGAA 1625

QY 1557 GCAGAAAGTTGAAAAACAATAGACTCAAGAAAGAAAAACAACAGTGATTTCCCTTCCTCA 1616

Db 1626 GCAGAAAGCTGAGAAAACAAATGACCTCAAGAAAAACAAACAAACAGCCATTTCCCTGCCGCA 1685

QY 1617 GATACTGGACTAAACAGCTTTCACCTGGTGTGGAGGGGCCACGAAGCAAGCCAAAGCTGA 1676

Db 1686 GATACTAGGACTGACAGCTTTCACCTGGTGTGGAGCGAGCAAAAAAGCAGTCTGAGGCTGA 1745

QY 1677 AGAACACATTTTAAAACTATGTGCCAATCTTGATGCAATTTACTATTAAAACTGTTAAAGA 1736

Db 1746 AAAACATATTTTAAATATATGTGCCAATCTTGATGCCTTTACCATTTAAACACAGTGAAGA 1805

QY 1737 AAACCTTGATCAACTGAAAAACCAATACAGAGGCCATGCAAGAAAGTTTGCCATTGCAGA 1796

Db 1806 GAATCTTGGTCAACTCAACACCAATAAAGGAACCATGCAAGAAATTTGTGATTGCTGA 1865

QY 1797 TGCAACCCAGAGAAGATCCATTTAAAGAGAACTTCTAGAAATAATGACAAGGATTCAAAC 1856

Db 1866 TGACACCAGAGAAAAATCCATTTAAAGAGAAACTTCTAGAAATTTATGGCAAGCATTCAGAC 1925

QY 1857 TTATTGTCAAATGAGTCCAAATGTCAGATTTTGAACTCAACCCCTATGAACAATGGGCCAT 1916

Db 1926 TTACTGCCAAAAAAGTCCAATGTCAGATTTTGGAACCCCAACATTTATGAGCAGTGGGCCAT 1985

QY 1917 TCAAAATGGAAAAAAGAGCTGCAAAAAAGGAAATCGCAAGAAACGTGTTTGTGCAGAAACA 1976

Db 1986 TCAAAATGGAAAAAAGAGCTGTTAAAGACGGAAATCGCAAGATCGCGTCTGTGCAGAAACA 2045

QY 1977 TTTGAGGAAGTACAAATGAGGCCCTACAAATTAATGACACAAATTCGAATGATAGATGCGTA 2036

Db 2046 TTTGAGGAAGTACAAACGAAGCCCTACAAATCAACGACACCGATCCGAATGATTGATGCATA 2105

QY 2037 TACTCATCTTGAACCTTCTATAATGAAGAGAAAGATAAGAAGTTTGCAGTCATAGAAGA 2096

Db 2106 TAGCCACCTGGAGGCATTTCTACACTGATGAGAAAGAAAGAGTTTCGAGTCTCT--CAA 2162

QY 2097 TGATAGTGATGAGGGTGGTATGATGAGTATTGTGATGGTGTGATGAAGATGAGGATGATTT 2156

Db 2163 TGACAGCGACAAGATGATGACGAGGCCAGCAGTTGCAATGACCAACTTAAGGGCGATGT 2222

QY 2157 AAAGAAACCTTTGAAACTGGATGAAACAGATAGATTTCTCATGACTTTTATTTTGTGAAAA 2216

Db 2223 AAAGAAATCTTTGAAACTGGACGAAACCGATGAAATTTCTCATGAAATTTGTTCTTTGATAA 2282

QY 2217 CAATAAAATGTTGAAAAGGCTGGCTGAAAAACCCAGAAATATGAAAAATGAAAAAGCTGACCAA 2276

Db 2283 CAAGAAAAATGTTGAAAAAACTAGCTGAAAAACCCAAAAATACGAGAATGAAAAAACTCATTA 2342

QY 2277 ATTAAGAAATACCATAAATGGAGCAATATATACTAGGACTGAGGAATCAGCACGAGGAATAAT 2336

Db 2343 ATTAAGAAAAACGATACTGGAACAATTCACAAGGTCTGAGGAGTCTCCGAGGAATTAAT 2402

QY 2337 CTTTACAAAAACACGACAGAGTGATATGCGCTTTTCCAGTGGATTACTGAAAAATGAATA 2396

Db 2403 TTTTACAAAAACACGACAGAGCACCTACGCACTTTCCAGTGGATCATGAAAAATGCAAA 2462

QY 2397 ATTTGCTGAAGTAGAGTCAAAAGCCCAACATCTGATTGGAGCTGGACACAGCAGTGAGTT 2456

Db 2463 GTTTGCGGAAGTTGGAGTCAAAGCGCATCACCTGATTGGCGCGGGGCACAGCAGTGAAGT 2522

QY 2457 CAAACCCCATGACACAGAAATGAACAAAAAGAAAGTCATTAGTAAATTTTCGCACTGGAAAAAT 2516

Db 2523 CAAAGCCCATGACTCAGACTGAACAAAAAGAAAGTCATTAGTAAATTTTCGCACTGGCGAAAT 2582

QY 2517 CAATCTGCTTATCGCTACCACAGTGGCGAGAAGAGTCTGGATATTTAAAGAATGTAACAT 2576

Db 2583 AAATCTGCTTATCGTACGACGGTGGCAGAGGAAGGCCCTGGATATCAAAGAGTGCAATAT 2642

QY 2577 TGTATCCGTTATGGTCTCGTCACCAATGAATAAGCCATGGTCCAGGCCCGTGGTTCGAGC 2636

Db 2643 TGTATTTCGTTATGGCCTTGTACGAAACAGATAGCCATGGTCCAGGCCCGGGGTTCGAGC 2702

QY 2637 CAGAGCTGATGAGAGCACCTACGTCCTGGTGTCTCACAGTGGTTCAGGAGTTATCGAACA 2696

Db 2703 CAGAGCTGATGAAAGCACGATATGTCTCTGGTCCACGAGAGTGGCTCAGGAGTTTACCGAACG 2762

QY 2697 TGAGACAGTTAATGATTTCCGAGAGAGATGATGTATAAAGCTATACAGATGCAAAAGTATAATGA 2756

Db 2763 GGAGATTGTTAATGATTTCCGAGAGAGATGATGTATAAAGCTATTAACCGTGTTCAAAA 2822

QY 2757 TATGAAACCCAGAGGAGTATGCTCATAAAGATTTTGGAAATTACAGATGCAAAAGTATAATGA 2816

Db 2823 CATGAAACCCAGAGGAGTATGCACATAAGATTTTGGAAATTGCAGGTGCAAAAGTATCCTGGA 2882

QY 2817 AAAAGAAATGAAAAACCAAGAGAAATATTGCCAAGCATTTACAAGAAATAAACCCATCACATA 2876

Db 2883 AAAAGAAATGAAAGTCAAAAGAGCATTTGCAAGCAATACAACGACAATCCATCGTTAAT 2942

QY 2877 AACTTTCTTTGCAAAAACTGCAGTGTGCTAGCCTGTTCTGGGGAAGATATCCATGTAAT 2936

Db 2943 AACACTTCTCTGCAAAAAATTTGTAGCATGCTGCTGCTCGGAGAGAAAAACATCCATGTCTAT 3002

QY 2937 TGAGAAAAATGCATCAGTCAATATGACCCCAAGAAATTCAGGAACTTTACATTTGTAAGAGA 2996

Db 3003 TGAGAAAGATGCATCATGTCAATATGACACCAAGAAATTCAGAGGACTCTACATTTGTAAGAGA 3062

QY 2997 AAACAAAGCACCTGCAAAAAGAGTGTGCCGACTATCAAAATAAATGGTGAATCATCTGCAA 3056

Db 3063 AAACAAAGCACCTGCAAAAAGAAAAATTTGCTGATTATCAGACCAATGGAGAGATTAATCTGCAA 3122

QY 3057 ATGTGCCCAGGCTTGGGAAACAATGATGGTGCAAAAAGGCTTAGATTTCCTTGTCTCAA 3116

Db 3123 GTGTGCCCAGGCTTGGGAAACAATGATGGTGCAAAAAGGTTTAGATTTCCTTGTCTTAA 3182

QY 3117 AATAAGGAATTTTGTAGTGGTTTTCAAAAAATAATTCAACAAAGAAACAATACAAAAAGTG 3176

Db 3183 AATAAGGAATTTTGTAGTCAATTTCAAAAAATAACTCACCGAAGAAACAGTACAAGAAGTG 3242

QY 3177 GGTAAGAAATACCTATCACATTTCCCAATCTTTGACTATTTCAGAAATGCTGTTTATTAGTGA 3236

Db 3243 GGTGGAATTCCTATCAGATTTCTGATCTTTGACTACTCAGAAATACCTGCTGATAGTGA 3302

QY 3237 TGAGGATTAGCACTTGATTGAAGATTTCTTTTAAATACTATCAGTTAAACATTTAATATG 3296

Db 3303 TGAAGATTAGCACTTGATTTCATGATTATTTTAAATACTGCCAACTCAACATTTGAATATG 3362

QY 3297 ATTATGATTAAATGTATTCAATTATGCTACAGAACTGACATAAGAATCAATAAA 3348

Db 3363 ATTTTAATTTGTTGTTTTCGCTACACTACTGAGCTAACTCTAGAATATGTAGA 3414

QY 1317 ATGATCGTGTATTGGATTAAAGTGGTGATACCCAACTGAAAAATATCATTTCCAGAAGT 1376
Db 1386 ATGGTATCGAATATTGGATTAAAGTGGCGATACCCAGCTGAAAAATATCATTTCCAGAAGT 1445
QY 1377 TGTCAAGTCTGTGATATTATTATCAGTACAGCTCAAACTCCTTGAAAACTCCCTCTTTAA 1436
Db 1446 TGTCAAATCTTACGATGTTATTATCAGCACTGCTCAAACTCCTTGAAAACTCCCTCTTTAA 1505
QY 1437 CTTGGAAAATGGAGAAAGATGCTGGTGTTCAAATTGTCAGACTTTTCCCTCATTTATGTA 1496
Db 1506 TCTGGAGAGTGGAGACGATGACGGTGTGCAGCTGTTCAGACTTCTCTCATTTATGTA 1565
QY 1497 TGAATGTCTACACCAACAAAGAACGAGTGTATAATAAACATCATGAGGCAATTATTGAT 1556
Db 1566 TGAGTGCCATCACACCAACAAGGAGCGAGTCTATAACAACATCATGAGACGATATTGAA 1625
QY 1557 GCAGAAGTTGAAAAACAATAGACTCAAGAAAGAAAAACAACAGTGATTTCCCTTCCTCA 1616
Db 1626 GCAGAAGCTGAGAAAACAATGACCTCAAGAAAAACAACAACAGCCATTTCCTGCGCA 1685
QY 1617 GATACTGGGACTPAACAGCTTTCACCTGGTGTGGAGGGCCACGAAGCAAGCAAGCTGA 1676
Db 1686 GATACTAGGACTGACAGCTTTCACCTGGTGTGGAGCGAGCCAAAAAGCAGTCTGAGGCTGA 1745
QY 1677 AGAACACATTTTAAAACTATGTGCCAATCTTGATGCAATTTACTATTAAAACTGTTAAAGA 1736
Db 1746 AAAACATATTTTAAATATATGTGCCAATCTTGATGCCCTTTACCATTTAAACAGTGAAGA 1805
QY 1737 AAACCTTGATCACTGAAAAACCAATACAGAGGCCATGCAAGAAGTTTGGCCATTCGAGA 1796
Db 1806 GAATCTTGGTCAACTCAACACCAAAATAAAGGAACCATGCAAGAAATTTGTGATTGCTGA 1865
QY 1797 TGCAACCAGAGAAGATCCATTTAAAGAGAAACTTCTAGAAATAATGACAAGGATTCAAAC 1856
Db 1866 TGACACCAGAGAAAAATCCATTTAAAGAGAAACTTCTAGAAATTTGGCAAGCAITTCAGAC 1925
QY 1857 TTATTGTCAAATGAGTCCAATGTCAGATTTTGGAACTCAACCCCTATGAACAATGGGCCAT 1916
Db 1926 TTACTGCCAAAAAAGTCCAATGTCAGATTTTGGAAACCAACATATGAGCAGTGGGCCAT 1985
QY 1917 TCAAAATGGAAAAAAGCTGCAAAAAAAGGAAATCGCAAAAGACGTGTTTGTGCAGAAC 1976
Db 1986 TCAAAATGGAAAAAAGCTGCTAAAGACGAAATCGCAAAAGATCGCGTCTGTGCAGAAC 2045
QY 1977 TTTGAGGAAGTACAAATGAGGCCCTCAAAATTAATGACACAATTCGAATGATAGATGCGTA 2036
Db 2046 TTTGAGGAAGTACAAACGAAGCCCTACAAATCAACGACACGATCCGAATGATTGATGCATA 2105
QY 2037 TACTCATCTTGAAACTTTCTATAATGAAGAGAAAGATAAGAAAGTTTGCAGTCAAGAGA 2096
Db 2106 TAGCCACCTGGAGGCAATCTACACTGATGAGAAAGAAAGAAAGTTGCGAGTCCT--CAA 2162
QY 2097 TGATAGTGATGAGGGTGGTGATGATGAGTATTGTGATGGTGTGATGAAGATGAGGATGATTT 2156
Db 2163 TGACAGCGACAAGAGTGATGACGAGGCCAGCAGTTGCAATGACCAACTTAAGGGCGATGT 2222
QY 2157 AAAGAAACCTTTGAAACTGGATGAAACAGATAGATTTCTCATGACTTTATTTTGAATA 2216
Db 2223 AAAGAAATCTTTGAAACTGGACGAAACGGATGAATTTCTCATGAATTTGTTCTTTGATA 2282
QY 2217 CAATAAAATGTTGAAAAGGCTGGCTGAAAACCCAGAAATATGAAAATGAAAAGCTGACCAA 2276
Db 2283 CAAGAAATGTTGAAAAAACTAGCTGAAAACCCCAAAATACGAGAATGAAAAACTCATTA 2342
QY 2277 ATTAAGAAATACCATAATGGACCAATATACTAGGACTGAGGAATCAGCACGAGGAATAAT 2336
Db 2343 ATTAAGAAACACGATACTGGAAACAATTCACAAGGTCTGAGGAGTCTCTCCGAGGAATTAT 2402
QY 2337 CTTTACAAAAACACGACAGATGTCATATGCGCTTTCCAGTGGATTACTGAAAATGAAA 2396
Db 2403 TTTTACAAAAACACGACAGAGCACCTACGCACTTTCCAGTGGATCATGAAAATGCAAA 2462
QY 2397 ATTTGCTGAAGTAGGAGTCAAGGCCCAACCATCTGATTGGAGCTGGACACAGCAGTGAGTT 2456

Db 2463 GTTTGCGGAAGTTGGAGTCAAGCGCATCACCTGATTGGCGGGGCACAGCAGTGAAGT 2522
QY 2457 CAAACCCATGACACAGAAATGAACAAAAAGAAAGTCAATTAGTAAATTTCCGACTGGAAAAAT 2516
Db 2523 CAAGCCCATGACTCAGACTGAACAAAAAGAAAGTCAATTAGTAAATTTCCGACTGGCGAAAT 2582
QY 2517 CAATCTGCTTATCGCTACCAACAGTGGCAGAAAGGTCTGGATATTTAAAGAAATGTAAAT 2576
Db 2583 AAATCTGCTTATCGCTACGACGGTGGCAGAGGAAGGCTTGGATATCAAAGAGTGAATAT 2642
QY 2577 TGTATCCGTTATGGTCTCGTCACCAATGAAATAGCCATGGTCCAGGCCCGTGGTCGAGC 2636
Db 2643 TGTATTTCGTTATGGCCTTGTCAAGAACGAGATAGCCATGGTCCAGGCCCGGTCGAGC 2702
QY 2637 CAGAGCTGATGAGAGCACCTACGTCCTGGTGTGCTCACAGTGGTTCAGGAGTTATCGAACA 2696
Db 2703 CAGAGCTGATGAAAGCACGTATGTCCTGGTCACCAGCAGTGGCTCAGGAGTTACCGAACG 2762
QY 2697 TGAGACAGTTAATGATTTCCGAGAGAAAGATGATGTATAAAGCTATACATTTGTGTTCAAAA 2756
Db 2763 GGAGATTGTTAATGATTTCCGAGAGAAAGATGATGTATAAAGCTATTAACCGTGTCAAAA 2822
QY 2757 TATGAAACCCAGAGAGATATGCTCTATAAGATTTTGGAAATACAGATGCAAAAGTATAATGA 2816
Db 2823 CATGAAACCCAGAGAGTATGCACATAAGATTTTGGAAATTGCAAGGTGCAAAAGTATCCTGGA 2882
QY 2817 AAAGAAAAATGAAACCAAGAGAAATATTGCCAAGCATTTACAAGAAATAACCCATCACTAAT 2876
Db 2883 AAAGAAAAATGAAAGTCAAAAGAAAGCATTTGCAAAAGCAATACAACGCAATCCATCGTTAAT 2942
QY 2877 AACTTTCTCTTGCAAAAAAGTGCAGTGTGTAGCTGTCTCTGGGGAAGATATCCATGTAAT 2936
Db 2943 AACACTTCTCTGCAAAAAATTTAGCATGTGTGTCTGCTCGGAGAAAAACATCCATGTCTAT 3002
QY 2937 TCAGAAAAATGCATCACTCAATATGACCCAGCAATTTCAAGGAATTTACATTTGTAAGAGA 2996
Db 3003 TGAGAAGATGCATCATGTCAATATGACACCAAGATTTCAAGGGAATTTACATTTGTAAGAGA 3062
QY 2997 AAACAAAGCACTGCAAAAGAAAGTGTGCCGACTATCAAAATAAATGCTGAAATCATCTGCAA 3056
Db 3063 AAACAAAGCACTGCAAAAGAAATTTGCTGATTTATCAGACCAATGGAGAGATTTATCTGCAA 3122
QY 3057 ATGTGGCAGGCTTGGGAAACAATGATGGTGCACAAAGGCTTAGATTTGCTCTCAA 3116
Db 3123 GTGTGGCCAGGCTTGGGAAACAATGATGGTGCACAAAGGTTTAGATTTGCTTCTTAA 3182
QY 3117 AATAAGGAATTTGTAGTGGTTTTCAAAAAATAATTCAACAAAGAAACAATACAAAAAGTG 3176
Db 3183 AATAAGGAATTTGTAGTCAATTTCAAAAAATAACTCACGGAAGAAACAGTACAAGAAAGTG 3242
QY 3177 GGTAGAATTACCTATCAATTTCCCAATCTTGACTATTTCAGAAATGCTGTTTATTAGTGA 3236
Db 3243 GGTGGAATTCCTATCAGATTTCTGATCTTGACTACTCAGAAATCTGCTTGTATAGTGA 3302
QY 3237 TGAGGATTAGCACTTGATTGAAGATTTCTTTTAAATACTATCAGTTAAACATTTAATATG 3296
Db 3303 TGAAGATTAGCACTTGATTGATGATTTATTTTAAATACTGCTGCTTGTATAGTGA 3362
QY 3297 ATTATGATTAAATGATTTCAATTATGCTACAGAACTGACATAAAGAAATCAATAAA 3348
Db 3363 ATTTTAATTTGTTTTCGCTACACTACTGAGCTAACTCTAGAATATGTAGA 3414

.RESULT 8
AY075132
LOCUS
DEFINITION Mus musculus HELICARD mRNA, complete cds.
ACCESSION AY075132
VERSION AY075132.1 GI:18698980
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

AY075132 3078 bp mRNA linear ROD 20-MAY-2002

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;
1 (bases 1 to 3078)
Kovacsovics, M., Martinon, F., Micheau, O., Bodmer, J. L., Hofmann, K.
and Tschopp, J.

TITLE Overexpression of Helicard, a CARD-Containing Helicase Cleaved during Apoptosis, Accelerates DNA Degradation

JOURNAL Curr. Biol. 12 (10), 838-843 (2002)

12015121
PUBMED

REFERENCE 2 (bases 1 to 3078)

AUTHORS Kovacsovics, M., Hofmann, K. and Tschopp, J.

TITLE Direct Submission

JOURNAL Submitted (24-JAN-2002) Biochemistry, University of Lausanne, ch
des Boveresses 155, Epalinges, VD 1066, Switzerland

FEATURES Location/Qualifiers

source

source	1. .3078
--------	----------

/organism="Mug mugcylus"

/organism="Mus musculus"

```
,{"accession": "AF011111",  
/mol_type="mRNA"
```

```
/db xref="taxon:10090"
```

1. 3078

/note="helicase: contains two CARD domains"

```
/codon start=1
```

```

/condition="HELIUM"
/product="HELIUM"

```

```

/product= 11811CARD
/product id="AA1.77205 1"

```

db_xref="GI:18698981" /
protein_id=AA17203.1

/translation="MSIVCSAEDSFRNLILFFRPRRLKMIQVPEVPLDHLIFLSAETKE
QILKKNINTCGNTSAEALLSTLEQGWPLGWTQMFEALEHSGNPLAARYVKPTLTDL
PSPSSETAHDECHLLTLOPTLVLDKLLINDVLDTCFEGLLTVEDRNRISAAGNSGN
ESGVRELLRRIVQKENWFSTFLDVLTRQNDALFQELTGGCCPEDPTDLANSHRDGP
AANECLLPAVDESITEAWNVDLILPEASCTDSVSTTESDTLAEGSVSCFDESILGS
NSNMGRDSGTMGSDSDSVIQTKRVSPEPELQRPYQMEVAQPALDGNKIITCLPTGS
GKTRVAVYITKDLHDKKQASESGKIVLVNKMVLAELQFRKEFPYLLKWYRIIGLS
GDTQLKISFPEVVKYDYVITTAQILLENLSLQKQKPAIPLQIILGTASPVGAAKKQSAEAKHIL
KEAVYNNIMRRYLVKQLRNNDLKQNKPAIPLQIILGTASPVGAAKKQSAEAKHIL
NICANDAFITIKVENLQQLKHQIKPECKKFVIADDTNPFGEKLEINWASIQTYC
QKSPMSDFGTQHYEQWAIQMEKKAADGNRKDRVCAEHLRKYNEALQINDTIRMIDAY
SHLETFTYDEKEKFAVLNDSKSDDEASSCNDQLKGDVKKSLKDETDFLWNLFFD
NKKWLKLAENPKYENEKLIKRLNTILEQFTRESSSRGIPTKTRQSTYALSQWIME
NAKFAEVGVKAHHILIGAGHSSEVKMPTQEQKEVISKFTFGEINLLIATTVAEGLDI
KECNIIVIRYGLVTNEIAMVQARGARADESTYVLVTSSSGVTEREIVNDFREKQMYK
AINRVQNMKPEEYAHKILEQVQSILEKKMKVKRSIAKQANDNPSLITLLCNKCSMLV
CSGENHVIKEMHVNMTPEFKGLYIVRENKALQKFADYQTNGBEII CKCGQAGWTMM
VHKGLDPLCLKIRNEVWFKNNSPKQYKKWVELPIRFDLDYSEYCLYSDD "

ORIGIN

Query Match 63.5%; Score 2135.4; DB 10; Length 3078;
Best Local Similarity 81.3%; Pred. NO. 0;
Matches 2504; Conservative 0; Mismatches 571; Indels 6; Gaps 2;

169 ATGTCGAATGGGTATTCCACAGACGAGAAATTTCGGCTATCTCATCTCGTGCTTCAGGGCC 228

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

229 AGGGTGAAAAATGTACATCCAGGTGGAGCCCTGTGCTGGACCTACCTGACCTTTCTCTGCTGCA 288

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

289 GAGGTGAAGGAGCAGATTTCAGAGGACAGTCGCCACCTCCGGAAACATGCAGGCAGTTGAA 348

7

349 CTGCTGCTGAGCACCTTGGAGAAGGGAGTCTGGCACCTTGGTTGGACTCGGCAATTTCGTG 408

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

Qv 409 GAGGCCCTCCGGAGAACCGGCAGCCCTCTGGCCGCCCGCTACATGAACCTGAGCTCAGC 468

[illegible]

Qv 469 GACTTGCCCTCTCCATCGTTTGAGAACGCTCATGATGAATATCTCCAACCTGCTGAACCTC 528

[illegible]

QY 1606 CCCCTCCTCAGATACTGGGACTAACAGCTTCACCTGGTGTGGAGGGCCACGAAGCAA 1665
Db 1441 CCCCTGCCGAGATACTAGGACTGACAGCTTCACCTGGTGTGGAGGAGCCAAAAGCAG 1500
QY 1666 GCCAAGCTGAAGAACACATTTTAAACCTATGIGCCAACTCTTGATGCACTTACTATTAAA 1725
Db 1501 TCTGAGCTGAAAACATATTTTAAATATATGIGCCAACTCTTGATGCTTTACCATTTAA 1560
QY 1726 ACTGTTAAAGAAACCTTGATCAACTGAAAAACCAAAATACAGGAGCCATGCAAGAACTTT 1785
Db 1561 ACAGTGAAGAGAATCTTGGTCAACTCAACACCAAAATAAGGAACCATGCAAGAAATTT 1620
QY 1786 GCCATTGCAGATGCAACCCAGAGAGATCCATTTAAAGAGAAACCTTCTAGAAATAATGACA 1845
Db 1621 GTGATTGCTGATGACACCAGAGAAAATCCATTTAAAGAGAAAACCTTCTAGAAAATTATGGCA 1680
QY 1846 AGGATTCAAACTTATTGTCAAATGAGTCCAATGTCAGATTTTGGAACTCAACCCCTATGAA 1905
Db 1681 AGCATTCAGACTTACTGCCAAAAAAGTCCAATGTCAGATTTTGGAACCCCAACATTATGAG 1740
QY 1906 CAATGGGCCATTCAAATGGAAAAAAGCTGCAAAAAAGGAAATCGCAAGAACGTGTT 1965
Db 1741 CAGTGGGCCATTCAAATGGAAAAAAGCTGTAAGACGGGAAATCGCAAGATCGCGTC 1800
QY 1966 TGTGCAGAACATTTGAGGAAGTACAATGAGGCCCTACAAATTAATGACACAATTCGAATG 2025
Db 1801 TGTGCAGAACATTTGAGGAAGTACAAGAACCCCTACAAATCAACGACACGATCCGAATG 1860
QY 2026 ATAGATGCGTATACTCATCTTGAAACTTTCTATAATGAAGAGAAAAGATAAGAACTTTGCA 2085
Db 1861 ATTGATGCATATAGCCACCTGGAGACATTTCTACACTGATGAGAAAAGAAAGATTCGCA 1920
QY 2086 GTCATAGAAGATGATGATGAGGGTGTGATGATGAGTATTGTGATGGTGAAGAT 2145
Db 1921 GTCTC- --AATGACAGGACAAGAGTGAAGGCCAGGCTGCAATGACCAACTT 1977
QY 2146 GAGGATGATTTAAAGAAACCTTTGAAACTGGATGAAACAGATAGATTTTCTCATGACTTTA 2205
Db 1978 AAGGCGATGTAAGAAATCTTTGAAACTGGACGAAACGGATGAATTTCTCATGAATTTG 2037
QY 2206 TTTTTTGAAACAATAAATGTTGAAAAGGCTGGCTGAAAACCCAGAAATATGAAAATGAA 2265
Db 2038 TTCTTTGATAACAAGAAAATGTTGAAAACCTAGCTGAAAACCCAAAATACGAGATGAA 2097
QY 2266 AAGCTGACCAAAATTAAGAAATACCATAATGGAGCAATATACTAGCACTGAGGAATCAGCA 2325
Db 2098 AAACCTCATTAATTAAGAAACACGATACCTGGAACAATTCACAAGGCTGAGGAGTCTCTCC 2157
QY 2326 CGAGGAATAATCTTTACAAAAACACGACAGAGTGCATATGCGCTTTCCAGTGGATTACT 2385
Db 2158 CGAGGAATATTTTTTACAAAAACACGACAGAGCACCTACGCACTTTCCAGTGGATCATG 2217
QY 2386 GAAATGAAAAAATTTGCTGAAGTAGGAGTCAAGCCCCACCATCTGATTTGGAGCTGGACAC 2445
Db 2218 GAAATGCAAAAGTTTTCGGAAGTTGGAGTCAAGGCGCATCACCTGATTTGGCGGGGCAC 2277
QY 2446 AGCAGTGAGTTCAAACCCCATGACACAGAAATGAACAAAAAGAGTCAATTAGTAAATTTCCG 2505
Db 2278 AGCAGTGAAGTCAAGCCCCATGACTCAGACTGAACAAAAAGAGTCAATTAGTAAATTTCCG 2337
QY 2506 ACTGGAATAATCAATCTGCTTATCGCTACCACAGTGGCAGAGAAAGGCTGGATATTTAA 2565
Db 2338 ACTGGCGAAATAAATCTGCTTATCGCTACGACGCTGGCAGAGGAGGCGCTGGATATCAA 2397
QY 2566 GAATGTAACATTTGTTATCCGTTTATGGTCTCGTCACCAATGAAATAGCCATGGTCCAGGCC 2625
Db 2398 GAGTGCAATATTGTTATTTCGTTTATGGCCTTGTACGAACGAGATAGCCATGGTCCAGGCC 2457
QY 2626 CGTGGTCGAGCCAGAGCTGATGAGAGCACCTACGTCCTGGTTGCTCACAGTGGTTTCAGGA 2685
Db 2458 CGGGGTCGAGCCAGAGCTGATGAAGCACGATGTCCTGGTCACCAGCAGTGGCTCAGGA 2517
QY 2686 GTTATCGAACATGAGACAGTTAATGATTTCCGAGAGAAAGATGATGTATAAAGCTATACAT 2745

Db 2518 GTTACCGAACGGAGATTGTTAATGATTTCCGAGAGAAGATGATGTATAAAGCTATTAAAC 2577
QY 2746 TGTGTTCAAAATATGAAACCAGAGGAGTATGCTCATAAGATTTTGAATTTACAGATGCAA 2805
Db 2578 CGTGTTCAAAACATGAAACCAGAGGAGTATGCACATAAAGATTTTGAATTTGCAGGTGCAA 2637
QY 2806 AGTATAATGAAAAAGAAAATGAAAAACCAAGAGAAAATATTGCCAAGCATTACAAGAATAAC 2865
Db 2638 AGTATCTGAAAAAGAAAATGAAAAGTCAAAAAGAACATTTGCAAAAGCAATACAACGACAAT 2697
QY 2866 CCATCACTAATAACTTTCTTTGCAAAAACCTGCAGTGTGCTAGCCTGTTCTGGGGAAGAT 2925
Db 2698 CCATCGTTAATAACACTTTCTTGCAAAAATTTGTAGCATGCTGGTCTGCTCGGGAGAAAAC 2757
QY 2926 ATCCATGTAATTGAGAAAATGCAATCAGTCATCAATATGACCCAGAAATTCAGGAACTTTAC 2985
Db 2758 ATCCATGTCATTGAGAAGATGCAATCATGTCAATATGACACAGAAATTCAGGGGACTCTAC 2817
QY 2986 ATTGTAAGAGAAAAACAAAGCACTGCAAAAAGAAAGTGTCCGCACTATCAAAATAAATGGTGAA 3045
Db 2818 ATTGTAAGAGAAAAACAAAGCACTGCAAAAAGAAATTTGCTGATTATCAGACCAATGGAGAG 2877
QY 3046 ATCATCTGCAAAATGTGGCCAGGCTTGGGGAACAATGATGGTGCACAAAGGCTTAGATTG 3105
Db 2878 ATTATCTGCAAGTGTGGCCAGGCTTGGGGAACAATGATGGTGCACAAAGGTTTAGATTG 2937
QY 3106 CCTTGTCTCAAAATAAGGAATTTTGTAGTGGTTTTTCAAAAATAATTCACAAAGAAACAA 3165
Db 2938 CCTTGTCTTAAAATAAGGAATTTTGTAGTCAATTTCAAAAATAACTCACCCGAGAAACAG 2997
QY 3166 TACAAAAAGTGGGTAGAAATTACCTATCACATTTCCCAATCTTGACTATTTCAGAAATGCTGT 3225
Db 2998 TACAAGAAAGTGGGTGGAATTTGCCATATCAGATTTCTGATCTTGACTACTCAGAAATACTGC 3057
QY 3226 TTATTTAGTGATGAGGATTAG 3246
Db 3058 TTGTATAGTGATGAAGATTAG 3078

RESULT 9

AX714144
LOCUS AX714144 1776 bp DNA linear PAT 15-APR-2003
DEFINITION Sequence 828 from Patent EP1293569.

ACCESSION AX714144

VERSION AX714144.1 GI:29889072

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS Isogai,T., Sugiyama,T., Otsuki,T., Wakamatsu,A., Sato,H., Ishii,S.,
Yamamoto,J.I., Isono,Y., Hio,Y., Otsuka,K., Nagai,K., Irie,R.,
Tamechika,I., Seki,N., Yoshikawa,T., Otsuka,M., Nagahari,K. and
Masuho,Y.

TITLE

Full-length cDNAs

Patent: Ep 1293569-A 828 19-MAR-2003;

Helix Research Institute (JP) ; Research Association for

Biotechnology (JP)

FEATURES

source

1..1776

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

ORIGIN

Query Match 52.5%; Score 1766; DB 6; Length 1776;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1769; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1592 ACAAACCACTGATTTCCCTTCTCAGATACTGGGACTAACAGCTTCACCTGGTTGGAG 1651
|||||

Db 2 ACAAACCAAGTGATTCCTCCCTTCCTCAGATACTGGGACTAACACAGCTTCACCTGGTGTGGAG 61

QY 1652 GGGCCACGAAGCAGCCAAAGCTGAAGAACACATTTTAAACTATGTGCCAATCTTGATG 1711

Db 62 GGGCCACGAAGCAGCCAAAGCTGAAGAACACATTTTAAACTATGTGCCAATCTTGATG 121

QY 1712 CATTCTATTAAAACTGTTTAAAGAAAACCTTGATCAACTGAAAAACCAAAATACAGGAGC 1771

Db 122 CATTCTACTATTAAAACTGTTTAAAGAAAACCTTGATCAACTGAAAAACCAAAATACAGGAGC 181

QY 1772 CATGCAAGAAGTTTGCCCATTTGCAGATGCAACCAAGAGAAGATCCATTTAAAGAGAACTTC 1831

Db 182 CATGCAAGAAGTTTGCCCATTTGCAGATGCAACCAAGAGAAGATCCATTTAAAGAGAACTTC 241

QY 1832 TAGAAAATAATGACAAAGGATTCAAAACCTTATTGTCAAATGAGTCCAATGTCCAGATTTTGGAA 1891

Db 242 TAGAAAATAATGACAAAGGATTCAAAACCTTATTGTCAAATGAGTCCAATGTCCAGATTTTGGAA 301

QY 1892 CTCAACCCCTATGAACAATGGGCCATTCAAATGGAAAAAAAAGCTGCAAAAAAGGAATC 1951

Db 302 CTCAACCCCTATGAACAATGGGCCATTCAAATGGAAAAAAAAGCTGCAAAAAAGGAATC 361

QY 1952 GCAAGAAACGTGTTGTGCAGAAACATTTGAGGAAAGTACAATGAGGCCCTACAAAATTAATG 2011

Db 362 GCAAGAAACGTGTTGTGCAGAAACATTTGAGGAAAGTACAATGAGGCCCTACAAAATTAATG 421

QY 2012 ACACAAATTCGAATGATAGATGCGTATACTCATCTTGAAACCTTTCTATAATGAAGAGAAAG 2071

Db 422 ACACAAATTCGAATGATAGATGCGTATACTCATCTTGAAACCTTTCTATAATGAAGAGAAAG 481

QY 2072 ATAAGAAGTTTGCGATCATAGAAGATGATAGTGAGGGTGGTGATGATGAGTATTGTG 2131

Db 482 ATAAGAAGTTTGCGATCATAGAAGATGATAGTGAGGGTGGTGATGATGAGTATTGTG 541

QY 2132 ATGGTGATGAAGATGAGGATGATTTTAAAGAAACCTTTGAAACTGGATGAAACAGATAGAT 2191

Db 542 ATGGTGATGAAGATGAGGATGATTTTAAAGAAACCTTTGAAACTGGATGAAACAGATAGAT 601

QY 2192 TTCTCATGACTTTATTTTTTGAAGAACATAAAATGTTGAAAGGCTGGCTGAAACCCAG 2251

Db 602 TTCTCATGACTTTATTTTTTGAAGAACATAAAATGTTGAAAGGCTGGCTGAAACCCAG 661

QY 2252 AATATGAAAATGAAAAGCTGACCAAATTAAGAAAATACCATAATGGAGCAATATACTAGGA 2311

Db 662 AATATGAAAATGAAAAGCTGACCAAATTAAGAAAATACCATAATGGAGCAATATACTAGGA 721

QY 2312 CTGAGGAATCAGCACGAGGAATAATCTTTACAAAAACACGACAGAGTGCATATGCGCTTT 2371

Db 722 CTGAGGAATCAGCACGAGGAATAATCTTTACAAAAACACGACAGAGTGCATATGCGCTTT 781

QY 2372 CCCAGTGGATTACTGAAAAATGAAAAATTTGCTGAAGTAGGAGTCAAAGCCCCACCATCTGA 2431

Db 782 CCCAGTGGATTACTGAAAAATGAAAAATTTGCTGAAGTAGGAGTCAAAGCCCCACCATCTGA 841

QY 2432 TTGGAGCTGGACACAGCAGTGAGTTCAAACCCCATGACACAGAAATGAACAAAAGAGTCA 2491

Db 842 TTGGAGCTGGACACAGCAGTGAGTTCAAACCCCATGACACAGAAATGAACAAAAGAGTCA 901

QY 2492 TTAGTAAATTTCCGACTGGAAAAATCAATCTGCTTATCGCTACCACAGTGGCAGAGAAG 2551

Db 902 TTAGTAAATTTCCGACTGGAAGAATAAATCTGCTTATCGCTACCACAGTGGCAGAGAAG 961

QY 2552 GTCTGGATATTAAAGAATGTAACATTGTTATCGTTATGTTCTCGTCACCAATGAATAG 2611

Db 962 GTCTGGATATTAAAGAATGTAACATTGTTATCGTTATGTTCTCGTCACCAATGAATAG 1021

QY 2612 CCATGGTCCAGSCCGTGGTCGAGCCAGAGCTGATGAGAGCACCTACGTCTGTTGCTC 2671

Db 1022 CCATGGTCCAGSCCGTGGTCGAGCCAGAGCTGATGAGAGCACCTACGTCTGTTGCTC 1081

QY 2672 ACAGTGGTTCAGAGTTATCGAAACATGAGACAGTTAATGATTTCCGAGAGAAGATGATG 2731

Db 1082 ACAGTGGTTCAGAGTTATCGAAACGTGAGACAGTTAATGATTTCCGAGAGAAGATGATG 1141

QY 2732 ATAAAGCTATACATTGTGTTCAAATAATGAAACCCAGAGAGTATGCTCATAGAATTTTGG 2791

Db 1142 ATAAAGCTATACATTGTGTTCAAATAATGAAACCCAGAGAGTATGCTCATAGAATTTTGG 1201

QY 2792 AATTACAGATGCAAAAGTATAATGGAAGAAAAATGAAAAACCAAGAGAAATATTGCCAAGC 2851

Db 1202 AATTACAGATGCAAAAGTATAATGGAAGAAAAATGAAAAACCAAGAGAAATATTGCCAAGC 1261

QY 2852 ATTACAAGAAATAACCCATCACTAATAACTTTCCTTTGCAAAAACTGCAGTGTGTAGCCT 2911

Db 1262 ATTACAAGAAATAACCCATCACTAATAACTTTCCTTTGCAAAAACTGCAGTGTGTAGCCT 1321

QY 2912 GTTCTGGGAAGATATCCATGTAATTGAGAAAAATGCATCACGTCAATATGACCCAGAAT 2971

Db 1322 GTTCTGGGAAGATATCCATGTAATTGAGAAAAATGCATCACGTCAATATGACCCAGAAT 1381

QY 2972 TCAAGGAACCTTTACATTGTAAGAGAAAAACAAAGCACTGCAAAAAGAGTGTGCCGACTATC 3031

Db 1382 TCAAGGAACCTTTACATTGTAAGAGAAAAACAAACACTGCAAAAAGAGTGTGCCGACTATC 1441

QY 3032 AAATAAATGGTGAAATCATCTGCAAAATGTGGCCAGGCTTGGGGAAACAATGATGTTGCACA 3091

Db 1442 AAATAAATGGTGAAATCATCTGCAAAATGTGGCCAGGCTTGGGGAAACAATGATGTTGCACA 1501

QY 3092 AAGGCTTAGATTGTCCTTGTCTCAAAATAAGGAATTTTGTAGTGGTTTTCAAATAAATT 3151

Db 1502 AAGGCTTAGATTGTCCTTGTCTCAAAATAAGGAATTTTGTAGTGGTTTTCAAATAAATT 1561

QY 3152 CAACAAAGAAAAACAATACAAAAAGTGGGTAGAAATTACCTATCACATTTCCCAATCTTGACT 3211

Db 1562 CAACAAAGAAAAACAATACAAAAAGTGGGTAGAAATTACCTATCACATTTCCCAATCTTGACT 1621

QY 3212 ATTCAGATGCTGTTTATTATTAGTGATGAGGATTAGCACTTGATTGAAGATTCTTTAAAA 3271

Db 1622 ATTCAGATGCTGTTTATTATTAGTGATGAGGATTAGCACTTGATTGAAGATTCTTTAAAA 1681

QY 3272 TACTATCAGTTAAACATTTAATATGATTATGATTAAATGATTTCATTATGCTACAGAACTG 3331

Db 1682 TACTATCAGTTAAACATTTAATATGATTATGATTAAATGATTTCATTATGCTACAGAACTG 1741

QY 3332 ACATAAGAATCAATAAAATGATTGTTTACTCTG 3365

Db 1742 ACATAAGAATCAATAAAATGATTGTTTACTCTG 1775

RESULT 10

AK056293

LOCUS

DEFINITION

AK056293 1776 bp mRNA linear PRI 30-JAN-2004

Homo sapiens cDNA FLJ31731 fis, clone NT2RI2006855, weakly similar to Homo sapiens RNA helicase (RIG-I) mRNA.

ACCESSION

AK056293

VERSION

AK056293.1 GI:16551653

KEYWORDS

oligo capping; fis (full insert sequence).

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1

AUTHORS

Ota,T., Suzuki,Y., Nishikawa,T., Otsuki,T., Sugiyama,T., Irie,R., Wakamatsu,A., Hayashi,K., Sato,H., Nagai,K., Kimura,K., Makita,H., Sekine,M., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T., Ishii,S., Yamamoto,J., Saito,K., Kawai,Y., Isono,Y., Nakamura,Y., Nagahari,K., Murakami,K., Yasuda,T., Iwayanagi,T., Wagatsuma,M., Shiratori,A., Sudo,H., Hosoiri,T., Kaku,Y., Kodaira,H., Kondo,H., Sugawara,M., Takahashi,M., Kanda,K., Yokoi,T., Furuya,T., Kikkawa,E., Omura,Y., Abe,K., Kamihara,K., Katsuta,N., Sato,K., Tanikawa,M., Yamazaki,M., Ninomiya,K., Ishibashi,T., Yamashita,H., Murakawa,K., Fujimori,K., Tanai,H., Kimata,M., Watanabe,M., Hiraoka,S., Chiba,Y., Ishida,S., Ono,Y., Takiguchi,S., Watanabe,S., Yosida,M., Hotuta,T., Kusano,J., Kanehori,K., Takahashi-Fujii,A., Hara,H., Tanase,T., Nomura,Y., Togiya,S., Komai,F., Hara,R., Takeuchi,K., Arita,M., Imose,N., Musashino,K., Yuuki,H., Oshima,A.,

Sasaki,N., Aotsuka,S., Yoshikawa,Y., Matsunawa,H., Ichihara,T.,
Shiohata,N., Sano,S., Moriya,S., Momiyama,H., Sato,N., Takami,S.,
Terashima,Y., Suzuki,O., Nakagawa,S., Senoh,A., Mizoguchi,H.,
Goto,Y., Shimizu,F., Wakebe,H., Hishigaki,H., Watanabe,T.,
Sugiyama,A., Takemoto,M., Kawakami,B., Yamazaki,M., Watanabe,K.,
Kumagai,A., Itakura,S., Fukuzumi,Y., Fujimori,Y., Komiyama,M.,
Tashiro,H., Tanigami,A., Fujiwara,T., Ono,T., Yamada,K., Fujii,Y.,
Ozaki,K., Hirao,M., Ohmori,Y., Kawabata,A., Hikiji,T., Kobatake,N.,
Inagaki,H., Ikema,Y., Okamoto,S., Okitani,R., Kawakami,T.,
Noguchi,S., Itoh,T., Shigeta,K., Senba,T., Matsumura,K.,
Nakajima,Y., Mizuno,T., Morinaga,M., Sasaki,M., Togashi,T.,
Oyama,M., Hata,H., Watanabe,M., Komatsu,T., Mizushima-Sugano,J.,
Sato,T., Shirai,Y., Takahashi,Y., Nakagawa,K., Okumura,K.,
Nagase,T., Nomura,N., Kikuchi,H., Masuho,Y., Yamashita,R.,
Nakai,K., Yada,T., Nakamura,Y., Ohara,O., Isogai,T. and Sugano,S.
Complete sequencing and characterization of 21,243 full-length
human cDNAs
Nat. Genet. 36 (1), 40-45 (2004)
14702039

TITLE

JOURNAL
PUBMED
REFERENCE
AUTHORS

Ishibashi,T., Kanehori,K., Yosida,M., Watanabe,S., Ishida,S.,
Ono,Y., Hotuta,T., Hiraoka,S., Murakawa,K., Takiguchi,S.,
Kusano,J., Watanabe,M., Fujimori,K., Tanai,H., Ishida,M.,
Yamashita,H., Chiba,Y., Sugiyama,T., Irie,R., Otsuki,T., Sato,H.,
Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y.,
Saito,K., Nishikawa,T., Kimura,K., Matsuo,K., Nakamura,Y.,
Sekine,M., Kikuchi,H., Kanda,K., Wagatsuma,M., Takahashi-Fujii,A.,
Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S.,
Nagahari,K., Masuho,Y., Nagai,K. and Isogai,T.
NEDO human cDNA sequencing project
Unpublished

TITLE

JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

3 (bases 1 to 1776)
Isogai,T., Otsuki,T. and Sugiyama,T.
Direct Submission

COMMENT

Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute,
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: RAB and
HRI.

FEATURES

source

Location/Qualifiers

1..1776
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2RI2006855"
/cell_line="NT2"
/cell_type="teratocarcinoma"
/clone_lib="NT2RI2"
/note="cloning vector: pME18SFL3-mRNA from NT2 neuronal
precursor cells treated 2-weeks mitotic inhibitor after
5-weeks retinoic acid (RA) induction.-majorly NT2 neuron"
250..1656
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB71141.1"
/db_xref="GI:16551654"
/translation="MTRIQTYCQSPMSDFGTQPYEQWAIQMEKAAKEGNRKERVCA
EHLRKYNEALQINDTIRMIDAYTHLETFFYNEEKDKKFAVIEDDSDEGGDDEYCDGDED
EDDLKKPLKLDETDRFLMTLFFENNMKLKRLAENPEYENEKLTKLNTIMEQYRTEE
SARGIIFTKTRQSAVALSQWITENEKEFAEVGVKHAHLNLIAGHSSEFKPMTQNEQKEVI
SKFRTGRINLLIATVAEGLDIKECNIVRYGLVTNLIAMVQARGRARADESTYVLV
AHSCSGVIERETVNDFREKMYKAIHCVQNMKPEEYAHKILELQMQSIMKKMTKRN
IAKHYNPNSLITFLCKNCSVLACSGEDIHVEKMHVNMTPFEKELYIVRENKTLQK
KCADYQINGEIIICKGQAGTMMVHKGLDPLCLKIRNFVVVFRKNNSTKKQYKKWVELP
ITFPNLDYSECCLFSDSD"

CDS

Query Match	52.5%;	Score 1766;	DB 9;	Length 1776;			
Best Local Similarity	99.7%;	Pred. No. 0;					
Matches 1769;	Conservative	0;	Mismatches	5;			
			Indels	0;			
			Gaps	0;			
QY	1592	ACAAACCAAGT	GATTC	CCCTTCCTCAGAT	ACTGGGACTAACAGCTT	CACCTTCACCTGGTGGAG	1651
Db	2	ACAAACCAAGT	GATTC	CCCTTCCTCAGAT	ACTGGGACTAACAGCTT	CACCTTCACCTGGTGGAG	61
QY	1652	GGCCACGAAG	CAAGCCAAAGCT	GAAGAACACATTTTAA	AACTATGTGCCAATCTTGATG	1711	
Db	62	GGCCACGAAG	CAAGCCAAAGCT	GAAGAACACATTTTAA	AACTATGTGCCAATCTTGATG	121	
QY	1712	CATTTACTAT	TAAAACTGTTAAAGAAA	AACTTGATCAACTG	AAAAAACC	CAATACAGGAGC	1771
Db	122	CATTTACTAT	TAAAACTGTTAAAGAAA	AACTTGATCAACTG	AAAAAACC	CAATACAGGAGC	181
QY	1772	CATGCAAGA	AGTTTGCCATTG	CAGATGCAAC	CCAGAGAAGATCCATTTAA	AGAGAACTTC	1831
Db	182	CATGCAAGA	AGTTTGCCATTG	CAGATGCAAC	CCAGAGAAGATCCATTTAA	AGAGAACTTC	241
QY	1832	TAGAAATA	TGACAAGGATT	CAAAACCTTAT	TGTCAATGAGTCCA	ATGTCAAGATTTTGGAA	1891
Db	242	TAGAAATA	TGACAAGGATT	CAAAACCTTAT	TGTCAATGAGTCCA	ATGTCAAGATTTTGGAA	301
QY	1892	CTCAACCC	TATGAACAATGGGCCAT	TCAAATGGA	AAAAAAGCTGCA	AAAAAAGGAAATC	1951
Db	302	CTCAACCC	TATGAACAATGGGCCAT	TCAAATGGA	AAAAAAGCTGCA	AAAAAAGGAAATC	361
QY	1952	GCAAAGAA	CGTGTGTCAGAA	CAATTTGAGGA	AGTACAATGAGG	CCCTACAAATTAATG	2011
Db	362	GCAAAGAA	CGTGTGTCAGAA	CAATTTGAGGA	AGTACAATGAGG	CCCTACAAATTAATG	421
QY	2012	ACACAAT	TCGAATGATAGAT	GCGTATAC	TATCATCTTGA	AACTTCTATAAGAGAGAAAG	2071
Db	422	ACACAAT	TCGAATGATAGAT	GCGTATAC	TATCATCTTGA	AACTTCTATAAGAGAGAAAG	481
QY	2072	ATAAGAA	AGTTTGAGTCATAGA	AGATGATAGT	GAGGGTGGT	GATGATGAGTATTGTG	2131
Db	482	ATAAGAA	AGTTTGAGTCATAGA	AGATGATAGT	GAGGGTGGT	GATGATGAGTATTGTG	541
QY	2132	ATGGTGA	TGAAGATGAGGAT	GATTTAAAGAA	AACTTTTGA	AACTGGATGAAACAGATAGAT	2191
Db	542	ATGGTGA	TGAAGATGAGGAT	GATTTAAAGAA	AACTTTTGA	AACTGGATGAAACAGATAGAT	601
QY	2192	TTCTCAT	GACTTTATTTT	TGAAAA	CAATAAAATGTTG	AAAAGCGTGGCTGAAAACCCAG	2251
Db	602	TTCTCAT	GACTTTATTTT	TGAAAA	CAATAAAATGTTG	AAAAGCGTGGCTGAAAACCCAG	661
QY	2252	AATATGA	AAATGAAAAGCT	GACCAAA	ATTAAAGAA	TACCATAATCGGAGCAATATACTAGGA	2311
Db	662	AATATGA	AAATGAAAAGCT	GACCAAA	ATTAAAGAA	TACCATAATCGGAGCAATATACTAGGA	721
QY	2312	CTGAGGA	ATCAGCACGAGGA	ATAATCTTTACA	AAAAAACACGACAG	AGTGCATATGCGCTTT	2371
Db	722	CTGAGGA	ATCAGCACGAGGA	ATAATCTTTACA	AAAAAACACGACAG	AGTGCATATGCGCTTT	781
QY	2372	CCCAGT	GGATTACTGAAA	ATGAAAAATTTGCT	CAAGTAGGAGTCA	AAAGCCCACCATCTGA	2431
Db	782	CCCAGT	GGATTACTGAAA	ATGAAAAATTTGCT	CAAGTAGGAGTCA	AAAGCCCACCATCTGA	841
QY	2432	TTGGAG	CTGGACACAGCAG	CTGAGTTCA	AAACCCATGACACAG	AAATGAACAAAAAGAGTCA	2491
Db	842	TTGGAG	CTGGACACAGCAG	CTGAGTTCA	AAACCCATGACACAG	AAATGAACAAAAAGAGTCA	901
QY	2492	TTAGTAA	ATTTTCGCACTGG	AAAAATCAATCT	GTCTTATCGTAC	CCACAGTGGCAGAAAG	2551
Db	902	TTAGTAA	ATTTTCGCACTGG	AAAAATCAATCT	GTCTTATCGTAC	CCACAGTGGCAGAAAG	961
QY	2552	GTCTGA	TATTAAAGAA	TGTAACATTG	TATCCGTTAT	CGTCTCGTCACCAATGAAATAG	2611
Db	962	GTCTGA	TATTAAAGAA	TGTAACATTG	TATCCGTTAT	CGTCTCGTCACCAATGAAATAG	1021
QY	2612	CCATGG	TCCAGGCCCCGT	GTTCGAGCCAG	AGCTGATGAGAG	CACCTACGTCCTGTTGCTC	2671

ORIGIN

Db 814 GAAGGGAAGAAATATCATCATCTGCCTCCCTACAGGAGTGGAAAAACAGAGTGGCTGTT 873
QY 1189 TACATTGCCAAGGATCATTAGACAAGAAAGAAAGCATCTGAGCCTGGAAAAAGTTATA 1248
Db 874 TACATTGCCAAGGATCATTAGACAAGAAAGAAAGCATCTGAGCCTGGAAAAAGTTATA 933
QY 1249 GTTCTTGTCAAATAAGGTACTGCTAGTTGAACAGCTCTCCGCAAGGAGTTCCAAACCATTT 1308
Db 934 GTTCTTGTCAAATAAGGTACTGCTAGTTGAACAGCTCTCCGCAAGGAGTTCCAAACCATTT 993
QY 1309 TTGAAGAAATGGTATCGTGTATTGGATTAAAGTGGTGATACCCAACTGAAAAATATCATTT 1368
Db 994 TTGAAGAAATGGTATCGTGTATTGGATTAAAGTGGTGATACCCAACTGAAAAATATCATTT 1053
QY 1369 CCAGAAAGTTGTCAAGTCCCTGTGATATTATTCAGTACAGCTCAAACTCCCTTGAAAACTCC 1428
Db 1054 CCAGAAAGTTGTCAAGTCCCTGTGATATTATTCAGTACAGCTCAAACTCCCTTGAAAACTCC 1113
QY 1429 CTCTTAAACTTGGAAAAATGGAGAAGATGCTGGTTCATTTGTCAGACTTTTCCCTCATTT 1488
Db 1114 CTCTTAAACTTGGAAAAATGGAGAAGATGCTGGTTCATTTGTCAGACTTTTCCCTCATTT 1173
QY 1489 ATCATTGATGAATGTCATCACACCAACAAAGAACGCTGTATATAATACATCATGAGGCAT 1548
Db 1174 ATCATTGATGAATGTCATCACACCAACAAAGAACGCTGTATATAATACATCATGAGGCAT 1233
QY 1549 TATTTGATGCAGAAGTTGAAAAACAATAGACTCAAGAAAGAAAAACAAACCGATGATTTCCC 1608
Db 1234 TATTTGATGCAGAAGTTGAAAAACAATAGACTCAAGAAAGAAAAACAAACCGATGATTTCCC 1293
QY 1609 CTTCTCAGATACTGGGACTAACAGCTTACCTGGTGTGGAGGGGCCACGAGCAAGCC 1668
Db 1294 CTTCTCAGATACTGGGACTAACAGCTTACCTGGTGTGGAGGGGCCACGAGCAAGCC 1353
QY 1669 AAAGCTGAAGAACACATTTTAAACCTATGTCCTAATCTTGATGCTATTACTATTAAACT 1728
Db 1354 AAAGCTGAAGAACACATTTTAAACCTATGTCCTAATCTTGATGCTATTACTATTAAACT 1413
QY 1729 GTTAAAGAAAACTTGATCAACTGAAAAACCAAAATACAGGAGCCATGCAAGAGTTGGCC 1788
Db 1414 GTTAAAGAAAACTTGATCAACTGAAAAACCAAAATACAGGAGCCATGCAAGAGTTGGCC 1473
QY 1789 ATTGCAGATGCAACACGAGAGATCCATTTAAAGAGAACTTCTAGAAAAATATGACAAGG 1848
Db 1474 ATTGCAGATGCAACACGAGAGATCCATTTAAAGAGAACTTCTAGAAAAATATGACAAGG 1533
QY 1849 ATTCAAACTTATTGTCAATGAGTCCAAATGTGATGATTTTGGAACTCAACCCCTATGACAA 1908
Db 1534 ATTCAAACTTATTGTCAATGAGTCCAAATGTGATGATTTTGGAACTCAACCCCTATGACAA 1593
QY 1909 TGGGCCATTCAAATGGAAAAAAAAGCTGCAAAAAAGGAAATCGCAAGAACGTGTTTGT 1968
Db 1594 TGGGCCATTCAAATGGAAAAAAAAGCTGCAAAAAAGGAAATCGCAAGAACGTGTTTGT 1617
QY 1969 GCAGAACATTTGAGGAAGTACAATGAGGCCCTACAAATTAATGACACAAATTCGAATGATA 2028
Db 1618 ----- 1617
QY 2029 GATGCGTATACATCTTGAAACTTTCTATAATGAAGAGAAAGATAAGAGTTTGCAGTC 2088
Db 1618 ----- 1617
QY 2089 ATAGAAGATGATAGTATGAGGGTGGTATGATGAGTATTGTGATGGTATGAAGATGAG 2148
Db 1618 ----- 1617
QY 2149 GATGATTTAAAGAAACCTTTGAAACTGGATGAAACAGATAGATTTCTCATGACTTTATTT 2208
Db 1618 ----- 1617
QY 2209 TTTGAAACCAATAAAAATGTTGAAAAGGCTGGCTGAAAACCCAGAAATATGAAAATGAAAAG 2268

Db 1618 ---GAAAAACAATAAATGTTGAAAAGGCTGGCTGAAAAACCCAGATATGAAAAATGAAAAG 1674
QY 2269 CTGACCAAAATTAAGAAATACCATTAATGGAGCAATATATACTAGGACTGAGGAATCAGCACGA 2328
Db 1675 CTGACCAAAATTAAGAAATACCATTAATGGAGCAATATATACTAGGACTGAGGAATCAGCACGA 1734
QY 2329 GGAATAATCTTTTACAAAAACACGACAGAGTGCATATGCGCTTTCCAGTGGATTACTGAA 2388
Db 1735 GGAATAATCTTTTACAAAAACACGACAGAGTGCATATGCGCTTTCCAGTGGATTACTGAA 1794
QY 2389 AATGAAAAATTTGCTGAAGTAGGAGTCAAAGCCACCATCTGATTGGAGCTGGACACAGC 2448
Db 1795 AATGAAAAATTTGCTGAAGTAGGAGTCAAAGCCACCATCTGATTGGAGCTGGACACAGC 1854
QY 2449 AGTGAGTTCAAAACCCATGACACAGAAATGAACAAAAAGAGTCTAGTAAATTTCCGCACT 2508
Db 1855 AGTGAGTTCAAAACCCATGACACAGAAATGAACAAAAAGAGTCTAGTAAATTTCCGCACT 1914
QY 2509 GAAAAAATCAATCTGCTTATCGCTACACAGTGGCAGAAAGGCTCTGATATTTAAAGAA 2568
Db 1915 GAAAAAATCAATCTGCTTATCGCTACACAGTGGCAGAAAGGCTCTGATATTTAAAGAA 1974
QY 2569 TGTAAACATTGTTATCCGTTATGCTCTCGTCACCAATGAATAGCCATGGTCCAG 2622
Db 1975 TGTAAACATTGTTATCCGTTATGCTCTCGTCACCAATGAATAGCCATGGTCCAG 2028

RESULT 12
AX300841
LOCUS AX300841 1443 bp DNA linear PAT 30-NOV-2001
DEFINITION Sequence 10 from Patent WO0185955.
ACCESSION AX300841
VERSION AX300841.1 GI:17382119
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Bahr, G., Cocude, C. and Capron, A.
TITLE Rh16 polypeptides and its fragments and polynucleotides encoding
said polypeptides and therapeutic uses
JOURNAL Patent: WO 0185955-A 10 15-NOV-2001;
Istac (FR) ; INSTITUT PASTEUR DE LILLE (FR)
FEATURES
source Location/Qualifiers
1 .1443
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 42.7%; Score 1435.6; DB 6; Length 1443;
Best Local Similarity 99.7%; Pred. No. 1.3e-286;
Matches 1438; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 687 GAAAGAAAACTGGTTCTCTGCAATTTCTGAATGTTCTCGTCAAAACAGGAAACAATGAAC 746
Db 1 GAAAGAAAACTGGTTCTCTGCAATTTCTGAATGTTCTCGTCAAAACAGGAAACAATGAAC 60
QY 747 TGTCCTCAAGAGTTAACAGGCTCTGATTGCTCAGAAAGCAATGCAGAGATTGAGAATTTATC 806
Db 61 TGTCCTCAAGAGTTAACAGGCTCTGATTGCTCAGAAAGCAATGCAGAGATTGAGAATTTATC 120
QY 807 ACAAGTTGATGGTCCCTCAAGTGGAGAGCAACTTCTTCAACCAAGTTCAGCCAAATCT 866
Db 121 ACAAGTTGATGGTCCCTCAAGTGGAGAGCAACTTCTTCAACCAAGTTCAGCCAAATCT 180
QY 867 GGAGAAAGGAGGCTCTGGGCGCATGGAGAAATACTCATCAGAATCATCTTTTGCAGATTCTTC 926
Db 181 GGAGAAAGGAGGCTCTGGGCGCATGGAGAAATACTCATCAGAATCATCTTTTGCAGATTCTTC 240
QY 927 TGTAGTTTTCAGAAATCAGACACAAAGTTTGGCAGAAAGAGTGTGAGTGTGATGAAAG 986

Db 241 TGTAGTTTCAGAAATCAGACACAAAGTTTGGCAGAAAGTAGTGTCTAGCTGCTTAGATGAAAG 300

QY 987 TCTTGGACATAACAGCAACATGGGCAGTGATTCAGGCACCATGGGAAGTGATTCAGATGA 1046

Db 301 TCTTGGACATAACAGCAACATGGGCAGTGATTCAGGCACCATGGGAAGTGATTCAGATGA 360

QY 1047 AGAGAAATGTGGCAGCAAGAGCATCCCGGAGCCAGAACTCCAGCTCAGGCGCTTACCAAAT 1106

Db 361 AGAGAAATGTGGCAGCAAGAGCATCCCGGAGCCAGAACTCCAGCTCAGGCGCTTACCAAAT 420

QY 1107 GGAAGTTGCCAGCCAGCCTTGGAAAGGAAGAATATCATCTGCTCCCTACAGGGAG 1166

Db 421 GGAAGTTGCCAGCCAGCCTTGGAAAGGAAGAATATCATCTGCTCCCTACAGGGAG 480

QY 1167 TGGAAAAACCCAGAGTGGCTGTTTACATTGCCAAGGATCACTTAGACAAGAGAAAAAGC 1226

Db 481 TGGAAAAACCCAGAGTGGCTGTTTACATTGCCAAGGATCACTTAGACAAGAGAAAAAGC 540

QY 1227 ATCTGAGCCTGGAAAAAGTTATAGTCTTGTGTCAAATAAGGTACTGCTAGTTGAACAGCTCTT 1286

Db 541 ATCTGAGCCTGGAAAAAGTTATAGTCTTGTGTCAAATAAGGTACTGCTAGTTGAACAGCTCTT 600

QY 1287 CCGCAAGGAGTCCAAACCAATTTTGAAGAAAATGGTATCGTGTATTGGATTAAAGTGTGA 1346

Db 601 CCGCAAGGAGTCCAAACCAATTTTGAAGAAAATGGTATCGTGTATTGGATTAAAGTGTGA 660

QY 1347 TACCCAACTGAAAAATATCATTTCCAGAAAGTTGTCAAGTCTCTGTGATATTATTATCAGTAC 1406

Db 661 TACCCAACTGAAAAATATCATTTCCAGAAAGTTGTCAAGTCTCTGTGATATTATTATCAGTAC 720

QY 1407 AGCTCAAATCCTTGAAAAACCTCCCTCTTAAACTTGGAAAATGGAGAAGATGCTGGTGTTC 1466

Db 721 AGCTCAAATCCTTGAAAAACCTCCCTCTTAAACTTGGAAAATGGAGAAGATGCTGGTGTTC 780

QY 1467 ATTGTCAGACTTTTCCCTCATTTATCATTTGATGAATGTCAATCACACCAACAAAGACAGT 1526

Db 781 ATTGTCAGACTTTTCCCTCATTTATCATTTGATGAATGTCAATCACACCAACAAAGACAGT 840

QY 1527 GTATAATAACATCATGAGGCATTATTGATGCAGAAAGTTGAAAAACAATAGACTCAAGAA 1586

Db 841 GTATAATAACATCATGAGGCATTATTGATGCAGAAAGTTGAAAAACAATAGACTCAAGAA 900

QY 1587 AGAAAAACAACCAAGTGATTCCTCTCCTCAGATACTGGGACTAACAGCTTCACCTGGTGT 1646

Db 901 AGAAAAACAACCAAGTGATTCCTCTCCTCAGATACTGGGACTAACAGCTTCACCTGGTGT 960

QY 1647 TGGAGGGCCGCAAGCAAGCCAAAGCTGAAGAACACATTTTAAACTATGTGCCAATCT 1706

Db 961 TGGAGGGCCGCAAGCAAGCCAAAGCTGAAGAACACATTTTAAACTATGTGCCAATCT 1020

QY 1707 TGATGCATTTACTATTAAAACTGTTAAAGAAAACCTTGATCAACTGAAAAACCAAAATACA 1766

Db 1021 TGATGCATTTACTATTAAAACTGTTAAAGAAAACCTTGATCAACTGAAAAACCAAAATACA 1080

QY 1767 GGAGCCATGCAAGAAAGTTTGCCATTGCAGATGCAACCCAGAGAAGATCCATTTAAAGAGAA 1826

Db 1081 GGAGCCATGCAAGAAAGTTTGCCATTGCAGATGCAACCCAGAGAAGATCCATTTAAAGAGAA 1140

QY 1827 ACTTCTAGAAATAATGACAAGGATTCAAACCTTATTGTCAAATGAGTCCAATGTCCAGATT 1886

Db 1141 ACTTCTAGAAATAATGACAAGGATTCAAACCTTATTGTCAAATGAGTCCAATGTCCAGATT 1200

QY 1887 TGGAACTCAACCTATGAACAATGGGCCATTCAAATGGAAAAAAAAGCTGCAAAAAAAGG 1946

Db 1201 TGGAACTCAACCTATGAACAATGGGCCATTCAAATGGAAAAAAAAGCTGCAAAAAAAGG 1260

QY 1947 AAATCGCAAGAACGTTGTTGTGCAGAACATTTGAGGAAGTACAATGAGGCCCTTACAAAT 2006

Db 1261 AAATCGCAAGAACGTTGTTGTGCAGAACATTTGAGGAAGTACAATGAGGCCCTTACAAAT 1320

QY 2007 TAATGACACAATTCGAATGATAGATGCGGTATACCTCATCTTGAAACTTTTCTATATGAAGA 2066

Db 1321 TAATGACACAATTCGAATGATAGATGCGGTATACCTCATCTTGAAACTTTTCTATATGAAGA 1380

QY 2067 GAAAGATAAGAACTTTTCAGTTCATAGAAAGATGATAGTATGATGAGGGTGGTGTATGATGAGTA 2126

Db 1381 GAAAGATAAGAACTTTTCAGTTCATAGAAAGATGATAGTATGATGAGGGTGGTGTATGATGAGTA 1440

QY 2127 TTG 2129

Db 1441 TTG 1443

BC025508 2468 bp mRNA linear ROD 20-JUL-2004
Mus musculus interferon induced with helicase C domain 1, mRNA
(CDNA clone IMAGE:5252257), complete cds.

BC025508
BC025508.1 GI:19343862

Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 2468)

Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettman,M., Madan,A., Young,A.C., Rodrigues,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.

Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

2 (bases 1 to 2468)

Strausberg,R.
Direct Submission
Submitted (06-MAR-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@hgri.nih.gov
Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,
Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C.,
McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W.,
Tsurgueon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
Young,A., Zhang,L.-H. and Green,E.D.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 56 Row: c Column: 12
This clone has the following problem: The cds is short compared to
the longest cds in the locus.

FEATURES

source

1. .2468
/organism="Mus musculus"
/mol_type="mRNA"
/strain="mix FVB/N, C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:5252257"
/tissue_type="Mammary tumor. WAP-TGF alpha model. 7 months
old, gross tissue."
/clone_lib="NCI CGAP_Mam5"
/lab_host="DH10B"
/note="Vector: PCMV-SPORT6"
1. .2468
/gene="Ifih1"
/note="synonyms: Helicard, Hlcd, MDA5, MDA-5"
/db_xref="LocusID:71586"
/db_xref="MGI:1918836"
444. .1847

CDS

/gene="Ifih1"
/codon_start=1
/product="Ifih1 protein"
/protein_id="AAH25508.1"
/db_xref="GI:19343863"
/db_xref="LocusID:71586"
/db_xref="MGI:1918836"
/translation="MASIQYQKSPMSDFGTQHYEQWAIQMEKKAADGNRKDRVCA
EHLRYNEALQINDTIRMIDAHSHLETFTYDEKFKFAVLNDSDESDDEASSCNDOLK
GDVKSILKDETFDLNLFNDKMKLKLAEKPKYENKLIKLNTILEQFTRSEES
SRGIIFTKRQSTYALSQWIMENAKFAEVGVKAHLIGAGHSSEVKPMTQTEQKEVIS
KFRTEINLLIATVAEEGLDIKECNIVIRYGLVTNEIAMVQARARADESTYVLVT
SSGSVTEREIVNDFREKMYKAINRVQNMKPREYAHKILELQVQSILEKKMKVKRSI
AKQYNDNPSLITLLCKNCSMLVCSGENIHVIEKMHVNMTPEFKGLYIVRENKALQKK
FADYQTNGEIICKGQAWGTMVHKGLDLPCLKIRFVNVFNKNSPKKYKKWVELPI
RFPDLDYSEYCLYSDYD"

ORIGIN

Query Match 41.7%; Score 1402; DB 10; Length 2468;
Best Local Similarity 82.9%; Pred. No. 1.2e-279;
Matches 1613; Conservative 0; Mismatches 330; Indels 3; Gaps 1;

QY	1403	GTACAGCTCAAAATCCTTGAAAACTCCCTCTTAAACTTGGAAATGGAGAAGATGCTGGTG	1462
Db	7	GTCCGGCTCAAAATCCTTGAAAACTCCCTCTTAAATCTGGAGAGTGGAGACGATGACGGTG	66
QY	1463	TTCAATTGTGACACTTTCCCTCATTTATCATTTGATGAATGTCAATCACACCAACAAGAAG	1522
Db	67	TGACAGCTGTGACACTTCTCTCTCATTTATCATTTGATGAGTGCCATCACACCAACAGGAGG	126
QY	1523	CAGTGATAATAACATCATGAGGCATTATTGATGCAGAAAGTTGAAAAACAATAGACTCA	1582
Db	127	CAGTCTATAACAACATCATGAGACGATATTGAAAGCAGAAAGCTGAGAAAACAATGACCTCA	186
QY	1583	AGAAAGAAAAACAAACACAGTGATTCCTTCAGATACCTGGGACTAACAGCTTCACCTG	1642
Db	187	AGAAACAAACAAACACAGCATTCCTTCCTGCGCAGATACCTAGGACTGACAGCTTCACCTG	246
QY	1643	GTGTTGGAGGGCCACGAAGCAAGCCAAAGCTGAAGAACACACATTTTAAAACTATGTGCCA	1702
Db	247	GTGTTGGAGCAGCCAAAGACAGTCTGAGGCTGAAAAACATATTTTAAATATATGTGCCA	306
QY	1703	ATCTTGATGCATTTACTATTAAACTGTTTAAAGAAAACTTGTATCAACTGAAAAACCCAAA	1762
Db	307	ATCTTGATGCCTTTACCAATTAAACAGTGAAGAGAATCTTGGTCAACTCAACACCCAAA	366
QY	1763	TACAGGACCATGCAAGAAGTTTGCCATTGTCAGATGCAACCCAGAGAGATCCATTTAAG	1822
Db	367	TAAAGAACCATGCAAGAATTTGTGATTGCTGATGACACCCAGAGAAAAATCCATTTAAG	426
QY	1823	AGAACTTCTAGAAATAATGCAAGGATTCAAACCTATTGTCAAATGAGTCCAATGTCTAG	1882

Db	427	AGAACTTCTAGAAATATTGGCAAGCATTCAGACTTACTGCCAAAAAGTCCAATGTCTAG	486
QY	1883	ATTTTGGAACTCAACCTATGAACAATGGGCCATTCAAATGGAAGAAAAAGCTGCAAAA	1942
Db	487	ATTTTGGAACTCAACCTATGAACAATGGGCCATTCAAATGGAAGAAAAAGCTGCTAAG	546
QY	1943	AAGGAATCGCAAGAACGTTTGTGCAGAACATTTGAGGAAGTACAAATGAGGCCCTAC	2002
Db	547	ACGGAATCGCAAGAACGTTTGTGCAGAACATTTGAGGAAGTACAAAGAACCTTAC	606
QY	2003	AAATTAATGACACAATTCGAATGATAGATGCGTATACTCATCTTGAACCTTCTATAATG	2062
Db	607	AAATCAACGACACGATCCGAATGATGATGCACATAGCCACCTGGAGACATTTACACTG	666
QY	2063	AAGAGAAAGATAAGAAAGTTTGCAGTCATAGAAGATCATAGTAGGGTGGTGTATGATG	2122
Db	667	ATGAGAAAGAAAAAGAGTTTCGAGTCTCT--CAATGACAGCGACGAGAGTGTATGACGAGG	723
QY	2123	AGTATTGTGATGGTGAAGATGAGGATGATTTTAAAGAAACCTTTGAAACTGGATGAAA	2182
Db	724	CCAGCAGTTGCAATGACCAACTTAAAGGGCGATGTAAAGAAATCTTTGAAACTGGACGAA	783
QY	2183	CAGATAGATTTCTCATGACTTTTATTTTGTGAAAAACAATAAAATGTTGAAAAGCTGGCTG	2242
Db	784	CGGATGAATTTCTCATGAATTTTGTCTTTGTATAACAAGAAAAATGTTGAAAAAACTAGCTG	843
QY	2243	AAAACCCAGAAATATGAAATGAAAGCTGACCAAAATTAAGAAATACCATATATGGAGCAAT	2302
Db	844	AAAACCCCAAATACGAGAAATGAAAAAATCTATTAAATTAAGAAAAACGATATCTGAAACAAT	903
QY	2303	ATACTAGGACTGAGGAATCAGCACGAGGAAATAATCTTTACAAAAACACGACAGATGCAT	2362
Db	904	TCACAAGGTCTGAGGAGTCTCTCCCGAGGAATATTATTTTACAAAAACACGACAGACACCT	963
QY	2363	ATCGGCTTTCCAGTGGATTACTGAAAAATGAAAAATTTGCTGAAGTAGGAGTCAAAGCCC	2422
Db	964	ACGCACCTTCCAGTGGATCATGGAATAATGCAAGTTTGGGAAAGTTGGAGTCAAGCGC	1023
QY	2423	ACCATCTGATTGGAGCTGGACACAGCAGTGAAGTTCAAACCCATGACACAGAAATGAACAAA	2482
Db	1024	ATCACCTGATTGGCGGGGCACAGCAGTGAAGTCAAGCCCATGACTCAGACTGAACAAA	1083
QY	2483	AAGAACTCATTAATTTTCGCACTGGAAAAATCAATCTGCTTATCGCTACACAGTGG	2542
Db	1084	AAGAACTCATTAATTTTCGCACTGGCGAAATAAATCTGCTTATCGCTACGACGGTGG	1143
QY	2543	CAGAAAGAGTCTGGATATTAAAGATGTAAACATTGTTATCCGTTATGGTCTCGTCAACA	2602
Db	1144	CAGAGGAAGGCTGGATATCAAGAGTGAATATTGTTATTCGTTATGGCTTGTACGA	1203
QY	2603	ATGAAATAGCCATGGTCCAGGCCCGTGGTCGAGCCAGAGCTGATGAGAGCACCTACGTC	2662
Db	1204	ACGAGATAGCCATGGTCCAGGCCCGGTTGCGAGCTGATGAAAGCACGATGTGTC	1263
QY	2663	TGGTTGCTCACAGTGGTTTCCAGGTTATCGAACATGAGACAGTTAATGATTTCCGAGAGA	2722
Db	1264	TGGTCAACAGCAGTGGTTCAGGAGTTTACCGAACGGGAGATTGTTAATGATTTCCGAGAGA	1323
QY	2723	AGATGATGTATAAAGCTATACATTGTTTCAAAATATGAAACCCAGAGGATGTCTCATATA	2782
Db	1324	AGATGATGTATAAAGCTATTAACCGTGTTCAAAAATGAAACCCAGAGGATGTGCACATA	1383
QY	2783	AGATTTTGGAAATTACAGATGCAAAAGTATAATGGAAGAAAAATGAAACCAAGAGAAATA	2842
Db	1384	AGATTTTGGAAATTGAGGTGCAAAAGTATCTCTGGAAGAAAAATGAAAGTCAAAAGAGCA	1443
QY	2843	TTGCCAAGCATTACAAGATAAACCCATCACTAATAAATTTCTTTTGCAAAACTGCAGTG	2902
Db	1444	TTGCAAGCAATAACAACGCAATCCATCGTTAATAAACACTTCTCTGCAAAAAATTTAGCA	1503
QY	2903	TGCTAGCTGTTCTGGGGAAGATATCCATGTAAATGAGAAAAATGCATCACGTCAATATGA	2962


```
Db 1504 TGCTGGTCTGCTCGGGAGAAACATCCATGTCATTGAGAAGATGCATCATGTCAATATGA 1563
QY 2963 CCCAGAAATCAAGGAACCTTTACATTTGTAAGAGAAAACAAAGCACTGCAAAAGAAAGTGTG 3022
Db 1564 CACCAGAAATCAAGGACTCTACATTTGTAAGAGAAAACAAAGCACTGCAAAAGAAATTTG 1623
QY 3023 CCGACTATCAAATAAATGGTGAAATCATCTGCAAAATGTGGCCAGGCTTGGGAAACAATGA 3082
Db 1624 CTGATTATCAGACCCAATGGAGAGATTATCTGCAAGTGTGGCCAGGCTTGGGAAACAATGA 1683
QY 3083 TGGTGCAAAAGGCTTAGATTGGCTTGTCTCAAAAATAAGGAAATTTTGTAGTGGTTTCA 3142
Db 1684 TGGTGCAAAAGGTTTAGATTGGCTTGTCTTAAATAAAGGAAATTTTGTAGTCAATTTCA 1743
QY 3143 AAAATAATTCAACAAAGAAACAATAACAAAAGTGGGTAGAAATTAACCTATCACATTTCCCA 3202
Db 1744 AAAATAACTCACCGAAGAAACAGTACAAGAGTGGTGGAATTTGCTATCAGATTTCCCTG 1803
QY 3203 ATCTTGACTATTTCAGAAATGCTGTTTATTTTAGTGATGAGGATTAGCACTTGATTGAAGATT 3262
Db 1804 ATCTTGACTACTCAGAAATACTGCTTGATATAGTGATGAAGATTAGCACTTGATTTCATGATT 1863
QY 3263 CTTTAAAATACTATCAGTTAAACATTTAAATATGATTATGATTAATGATTATTCATTATGCT 3322
Db 1864 ATTTTAAATAATTGCCAACTCAACATTTGAATATGATTTTAAATTGTTGTTTCGCTACACT 1923
QY 3323 ACAGAACTGACATAAGAATCAATAA 3348
Db 1924 ACTGAGCTAACTCTAGAATATGTAGA 1949

RESULT 14
LOCUS AR379503 1392 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 48 from patent US 6607879.
ACCESSION AR379503
VERSION AR379503.1 GI:40087137
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 1392)
Cocks,B.G., Stuart,S.G. and Seilhamer,J.J.
TITLE Compositions for the detection of blood cell and immunological response gene expression
JOURNAL Patent: US 6607879-A 48 19-AUG-2003;
FEATURES Location/Qualifiers
source 1..1392
/mol_type="genomic DNA"

ORIGIN
Query Match 38.0%; Score 1278.8; DB 6; Length 1392;
Best Local Similarity 94.1%; Pred. No. 3.8e-254;
Matches 1311; Conservative 0; Mismatches 77; Indels 5; Gaps 2;

QY 1976 ATTTGAGGAAGTACAATGAGGCCCTACAAATTAATGACACAAATTCGAATGATAGATGCGT 2035
Db 1 ATTTGAGGAAGTACAATGAGGCCCTACAAATTAATGACACAAATTCGAATGATAGATGCGT 60

QY 2036 ATACTCATCTTGAAACTTTCTATAATGAAGAGAAAGATAAGAAGTTTGCAGTCATAGAAG 2095
Db 61 ATACTCATCTTGAAACTTTCTATAATGAAGAGAAAGATAAGAAGTTTGCAGTCATANNNN 120

QY 2096 ATGATAGTGATGAGGGTGGTGATGATGAGTATTGTGATGGTGATGAAGATGAGGATGATT 2155
Db 121 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNT 180

QY 2156 TAAAGAAACCTTTGAAACTGGATGAACACAGATAGATTTCTCATGACTTTTATTTTGA 2215
Db 181 TAAAGAAACCTTTGAAACTGGATGAACACAGATAGATTTCTCATGACTTTTATTTTGA 240

QY 2216 ACAATAAAATGTTGAAAGGGCTGGCTGAAACCAGAAATATGAAAATGAAAAGCTGACCA 2275
```

```
Db 241 ACAATAAAATGTTGAAAGGGCTGGCTGAAAACCCAGAAATATGAAAATGAAAAGCTGACCA 300
QY 2276 AATTAAGAAATACCATAAATGGAGCAATATACTAGGACTGAGGAATCAGCACGAGGAATAA 2335
Db 301 AATTAAGAAATACCATAAATGGAGCAATATACTAGGACTGAGGAATCAGCACGAGGAATAA 360
QY 2336 TCTTTACAAAACACGACAGAGTGCATATGCGCTTTCCAGTGGATTACTGAAAATGAAA 2395
Db 361 TCTTTACAAAACACGACAGAGTGCATATGCGCTTTCCAGTGGATTACTGAAAATGAAA 420
QY 2396 AATTGCTGAAGTAGGAGTCAAAAGCCACCACATCTGTATTGGAGCTGGACACAGCAGTGAGT 2455
Db 421 AATTGCTGAAGTAGGAGTCAAAAGCCACCACATCTGTATTGGAGCTGGACACAGCAGTGAGT 480
QY 2456 TCAAAACCCATGACACAGAAATGAACAAAAGAAAGTCAATTAGTAAATTTCCGACTGGAAAAA 2515
Db 481 TCAAAACCCATGACACAGAAATGAACAAAAGAAAGTCAATTAGTAAATTTCCGACTGGAAAAA 540
QY 2516 TCAATCTGCTTATCGCTACCAACAGTGGCAGAAAGGCTCTGGATATTAAAGAAATGTAACA 2575
Db 541 TAAATCTGCTTATCGCTACCAACAGTGGCAGAAAGGCTCTGGATATTAAAGAAATGTAACA 600
QY 2576 TTGTTATCCGTTATGGTCTCGTCAACAATGAATAGCCATGGTCCAGGCCCGTGGTCGAG 2635
Db 601 TTGTTATCCGTTATGGTCTCGTCAACAATGAATAGCCATGGTCCAGGCCCGTGGTCGAG 660
QY 2636 CCAGAGCTGATGAGAGCACCTACGTCCTGTTGCTCACAGTGGTTCAGGAGTTATCGAAC 2695
Db 661 CCAGAGCTGATGAGAGCACCTACGTCCTGTTGCTCACAGTGGTTCAGGAGTTATCGAAC 720
QY 2696 ATGAGACAGTTAATGATTTCCGAGAGAAAGATGATGTATAAAGCTATACATTTGTTCAAA 2755
Db 721 GTGAGACAGTTAATGATTTCCGAGAGAAAGATGATGTATAAAGCTATACATTTGTTCAAA 780
QY 2756 ATATGAAACCCAGAGGAGTATGCTCATAAGATTTTGGAAATTACAGATGCAAAAGTATAATG 2815
Db 781 ATATGAAACCCAGAGGAGTATGCTCATAAGATTTTGGAAATTACAGATGCAAAAGTATAATG 840
QY 2816 AAAAGAAAATGAAAACCAAGAGAAATATTGCCAAGCATTTACAAGAATAACCCATCACTAA 2875
Db 841 AAAAGAAAATGAAAACCAAGAGAAATATTGCCAAGCATTTACAAGAATAACCCATCACTAA 900
QY 2876 TAACTTTCTTTGCAAAAACCTGCAGTGTGCTAGCCTGTCTCTGGGGAAGATATCCATGTAA 2935
Db 901 TAACTTTCTTTGCAAAAACCTGCAGTGTGCTAGCCTGTCTCTGGGGAAGATATCCATGTAA 960
QY 2936 TTGAGAAAATGCATCACGTCAATATGACCCAGAAATTCAGGAACTTTTACATTTGTAAGAG 2995
Db 961 TTGAGAAAATGCATCACGTCAATATGACCCAGAAATTCAGGAACTTTTACATTTGTAAGAG 1020
QY 2996 AAAACAAAGCACTGCAAAAGAAAGTGTGCCGACTATCAATAAATGGTGAATCATCTGCA 3055
Db 1021 AAAACAAAGCACTGCAAAAGAAAGTGTGCCGACTATCAATAAATGGTGAATCATCTGCA 1080
QY 3056 AATGTGCCAGGCTTGGGGAAACNAATGATGGTGCACAAAGGCTTAGATTTGCCCTTGCTCA 3115
Db 1081 AATGTGCCAGGCTTGGGGAAACNAATGATGGTGCACAAAGGCTTAGATTTGCCCTTGCTCA 1140
QY 3116 AAATAAGGAATTTTGTAGTGGTTTCAAAAATAATTCAACAAAGAAAACAATACAAAAGT 3175
Db 1141 AAATAAGGAATTTTGTAGTGGTTTCAAAAATAATTCAACAAAGAAAACAATACAAAAGT 1200
QY 3176 GGTAGAAATACCTATCACATTTCCCAATCTTGACTATTTCAGAAATGCTGTTTATTAGTG 3235
Db 1201 GGTAGAAATACCTATCACATTTCCCAATCTTGACTATTTCAGAAATGCTGTTTATTAGTG 1260
QY 3236 ATGAGGATTAGCACTTGATTGAAGATTCTTTTAAAAATACTATCAGTTAAACATTTAATAT 3295
Db 1261 ATGAGGATTAGCACTTGATTGAAGATTCTTTTAAAAATACTATCAGTTAAACATT--AATA 1318
QY 3296 GATTATGATTAAATGATTTCATTATATGCTACAGAACTGACATAAGAATC--AATAAATGA 3352
|||||
```

Db 1319 TGTTATGATTAATGTATTCAATTATGCTACAGAACTGACAATNGAGNCTCAATAAAAAATGA 1378

QY 3353 TTGTTTTACTCTG 3365

Db 1379 TTGTTTTACTCTG 1391

RESULT 15

AX300838

LOCUS AX300838 1284 bp DNA linear PAT 30-NOV-2001

DEFINITION Sequence 7 from Patent WO0185955.

ACCESSION AX300838

VERSION AX300838.1 GI:17382116

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1

AUTHORS Bahr,G., Cocude,C. and Capron,A.

TITLE Rh116 polypeptides and its fragments and polynucleotides encoding said polypeptides and therapeutic uses

JOURNAL Patent: WO 0185955-A 7 15-NOV-2001;

Istac (FR) ; INSTITUT PASTEUR DE LILLE (FR)

FEATURES

source 1. .1284

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

ORIGIN

Query Match 37.3%; Score 1254.6; DB 6; Length 1284;

Best Local Similarity 99.6%; Pred. No. 4e-249;

Matches 1257; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2103 TGATGAGGGTGGTGATGATGAGTATTTGTGATGGTGATGAAGATGAGGATGATTTAAAGAA 2162

Db 1 TGATGAGGGTGGTGATGATGAGTATTTGTGATGGTGATGAAGATGAGGATGATTTAAAGAA 60

QY 2163 ACCTTTGAAACTGGATGAAACAGATAGATTTCTCATGACTTTATTTTGTGAAACAAATAA 2222

Db 61 ACCTTTGAAACTGGATGAAACAGATAGATTTCTCATGACTTTATTTTGTGAAACAAATAA 120

QY 2223 AATGTTGAAAAGGCTGGCTGAAAACCCAGAAATATGAAAATGAAAAGCTGACCAAAATTAAG 2282

Db 121 AATGTTGAAAAGGCTGGCTGAAAACCCAGAAATATGAAAATGAAAAGCTGACCAAAATTAAG 180

QY 2283 AAATACCATAATGGAGCAATATATACTAGGACTGAGGAATCAGCACGAGGAATAATCTTTAC 2342

Db 181 AAATACCATAATGGAGCAATATATACTAGGACTGAGGAATCAGCACGAGGAATAATCTTTAC 240

QY 2343 AAAAAACGACAGAGTGCCATATCGGCTTTCCAGTGGATTACTGAAAATGAAAATTTGC 2402

Db 241 AAAAAACGACAGAGTGCCATATCGGCTTTCCAGTGGATTACTGAAAATGAAAATTTGC 300

QY 2403 TGAAGTAGGAGTCAAAAGCCCACTCTGATTGGAGCTGGACACAGCAGTGAGTTCAAACC 2462

Db 301 TGAAGTAGGAGTCAAAAGCCCACTCTGATTGGAGCTGGACACAGCAGTGAGTTCAAACC 360

QY 2463 CATGACACAGAATGAACAAAAGAGTCAATTAGTAATTTCCGACTGGAAAAATCAATCT 2522

Db 361 CATGACACAGAATGAACAAAAGAGTCAATTAGTAATTTCCGACTGGAAAAATCAATCT 420

QY 2523 GCTTATCGCTACCA CAGTGGCAGAGAAGGTCTGGATATTAAAGATGTAAACATTGTTAT 2582

Db 421 GCTTATCGCTACCA CAGTGGCAGAGAAGGTCTGGATATTAAAGATGTAAACATTGTTAT 480

QY 2583 CCGTTATGGTCTCGT CACCAATGAATAGCCATGGTCCAGGCCCGTGGTCGAGCCAGAGC 2642

Db 481 CCGTTATGGTCTCGT CACCAATGAATAGCCATGGTCCAGGCCCGTGGTCGAGCCAGAGC 540

QY 2643 T6ATGAGAGCACCTACGTCCTGGTTGCTCA CAGTGGTTGAGGAGTTATCGAACATGAGAC 2702

Db 541 TGATGAGAGCACCTACGTCCTGGTTGCTCACAGTGGTTTCAGGAGTTATCGAACGTGAGAC 600

QY 2703 AGTTAATGATTTCCGAGAGAAGATGATGTATAAAGCTATACATTGTGTTCAAAATATGAA 2762

Db 601 AGTTAATGATTTCCGAGAGAAGATGATGTATAAAGCTATACATTGTGTTCAAAATATGAA 660

QY 2763 ACCAGAGGAGTATGCTCATAAAGATTTTGGAAATTACAGATGCAAAAGTATAATGGAAGAA 2822

Db 661 ACCAGAGGAGTATGCTCATAAAGATTTTGGAAATTACAGATGCAAAAGTATAATGGAAGAA 720

QY 2823 AATGAAAACCAAGAGAAATATGCCCAAGCATTACAAGAATAACCCATCACTAATAACTTT 2882

Db 721 AATGAAAACCAAGAGAAATATGCCCAAGCATTACAAGAATAACCCATCACTAATAACTTT 780

QY 2883 CCTTTGCAAAAACTGCAGTGTGCTAGCCTGTTCTGGGAAGATATCCATGTAAATTGAGAA 2942

Db 781 CCTTTGCAAAAACTGCAGTGTGCTAGCCTGTTCTGGGAAGATATCCATGTAAATTGAGAA 840

QY 2943 AATGCATCACGTCAATATGACCCCAAGAAATTCAGAGAACTTTACATTTGTAAGAGAAAAACAA 3002

Db 841 AATGCATCACGTCAATATGACCCCAAGAAATTCAGAGAACTTTACATTTGTAAGAGAAAAACAA 900

QY 3003 AGCACTGCAAAAAGAGTGTGCCGACTATCAAAATAAATGGTGAATCATCTGCAAAATGTGG 3062

Db 901 AACACTGCAAAAAGAGTGTGCCGACTATCAAAATAAATGGTGAATCATCTGCAAAATGTGG 960

QY 3063 CCAGGCTTGGGGAACAATGATGTGCACAAAGGCTTAGATTTGCCCTTGTCTCAAAATAAG 3122

Db 961 CCAGGCTTGGGGAACAATGATGTGCACAAAGGCTTAGATTTGCCCTTGTCTCAAAATAAG 1020

QY 3123 GAATTTTGTAGTGGTTTTCAAAAAATAATTCAACAAAGAAACAATACAAAAAGTGGGTAGA 3182

Db 1021 GAATTTTGTAGTGGTTTTCAAAAAATAATTCAACAAAGAAACAATACAAAAAGTGGGTAGA 1080

QY 3183 ATTACCTATCACATTTCCCAATCTTGACTATTCAGAAATGCTGTTTATTTAGTGATGAGGA 3242

Db 1081 ATTACCTATCACATTTCCCAATCTTGACTATTCAGAAATGCTGTTTATTTAGTGATGAGGA 1140

QY 3243 TTAGCACTTGATTGGAAGATTTCTTTTAAATACTATCAGTTTAAACATTTTAATATGATTATG 3302

Db 1141 TTAGCACTTGATTGGAAGATTTCTTTTAAATACTATCAGTTTAAACATTTTAATATGATTATG 1200

QY 3303 ATTAATGTATTTCATTATGCTACAGAACTGACATAAGAATCAATAAAATGATTGTTTTACT 3362

Db 1201 ATTAATGTATTTCATTATGCTACAGAACTGACATAAGAATCAATAAAATGATTGTTTTACC 1260

QY 3363 CT 3364

Db 1261 NT 1262

Search completed: January 27, 2005, 15:18:48

Job time : 13661 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 27, 2005, 10:27:35 ; Search time 1441 Seconds
(without alignments)
12258.364 Million cell updates/sec

Title: US-09-515-363C-1
Perfect score: 3365
Sequence: 1 gcgcgcgcgcctgagagccc.....aaaatgattgttttactctg 3365

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues 8269772

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_23Sep04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Query Length	DB ID	Description
1	3365	100.0	3365	4 AAD17203	Aad17203 Human mel
2	3365	100.0	3380	12 ADJ74904	Adj74904 Marker ge
3	3353	99.6	3379	12 ADN04879	Adn04879 Antipsori
4	3339.2	99.2	3668	12 ADQ22237	Adq22237 Human sof
5	3338.8	99.2	3372	6 ABA04908	Aba04908 Human RNA
6	3316.2	98.5	3446	10 ADC30823	Adc30823 Human nov
7	3104	92.2	3131	4 AAD11170	Aad11170 Human mel
8	2202.4	65.5	3771	12 ADJ75813	Adj75813 Marker ge
9	1878	55.8	1967	4 AAS40960	Aas40960 cDNA enco
10	1766	52.5	1776	10 ADA53260	Ada53260 Human cod
11	1460.2	43.4	1557	4 AAI59285	Aai59285 Human pol
12	1435.6	42.7	1443	6 ABA04916	Aba04916 Human RNA
13	1316.6	39.1	1382	4 AAI61071	Aai61071 Human pol
14	1316.6	39.1	1382	10 ADC32572	Adc32572 Human nov
15	1278.8	38.0	1392	11 ADI30722	Adi30722 Human CDN
16	1254.6	37.3	1284	6 ABA04913	Aba04913 Human RNA
17	1130.8	33.6	1319	4 AAS41517	Aas41517 cDNA enco
18	803	23.9	1258	4 AAS01149	Aas01149 Interfero
19	803	23.9	1270	4 AAS01151	Aas01151 Interfero
20	738.8	22.0	956	5 AAS91688	Aas91688 DNA encod
21	663	19.7	3692	4 AAS01150	Aas01150 Interfero

22	663	19.7	3704	4 AAS01152	Aas01152 Interfero
23	577.2	17.2	609	4 AAH33745	Aah33745 Human col
24	577.2	17.2	609	6 ABQ54189	Abq54189 Human ova
25	556	16.5	6406	4 AAD17207	Aad17207 Human mel
26	554	16.5	1036	4 AAD17204	Aad17204 Human mel
C 27	507.2	15.1	595	4 ABA61000	Aba61000 Human foe
C 28	507.2	15.1	595	4 AAI40897	Aai40897 Probe #95
C 29	507.2	15.1	595	4 AAK35181	Aak35181 Human bon
C 30	507.2	15.1	595	4 AAK09292	Aak09292 Human bon
C 31	455.2	13.5	499	4 ABA73600	Aba73600 Human foe
C 32	455.2	13.5	499	4 AAI54039	Aai54039 Probe #22
C 33	455.2	13.5	499	4 AAK48209	Aak48209 Human bon
C 34	455.2	13.5	499	4 AAK22046	Aak22046 Human bra
C 35	453	13.5	453	12 ACH90906	Ach90906 Human gen
36	419.4	12.5	457	9 ACH25741	Ach25741 Human adu
37	376.4	11.2	458	3 AAA64601	Aaa64601 Nucleotid
38	371.2	11.0	392	5 AAF66485	Aaf66485 Novel hum
39	365.8	10.9	374	2 AAV89625	Aav89625 EST clone
40	356.8	10.6	387	5 AAF66499	Aaf66499 Novel hum
C 41	342	10.2	501	12 ACH77200	Ach77200 Human gen
42	331.4	9.8	341	4 AAL23623	Aal23623 Human bre
43	330.4	9.8	340	4 AAL14761	Aal14761 Human bre
44	285.4	8.5	301	2 AAQ90050	Aaq90050 Melanoma
45	281.4	8.4	292	4 AAI19913	Aai19913 Probe #98

ALIGNMENTS

RESULT 1

AAD17203

ID AAD17203 standard; cDNA; 3365 BP.

XX AAD17203;

DT 29-NOV-2001 (first entry)

XX Human melanoma differentiation associated (Mda)-5 cDNA.

KW Human; melanoma differentiation associated gene; Mda-5; interferon; IFN;
KW RNA helicase motif; caspase recruitment domain; CARD; therapy; melanoma;
KW neuroblastoma; astrocytoma; glioblastoma; multiforme; cancer; cervical;
KW breast; colon; prostate; osteosarcoma; chondrosarcoma; systemic toxicity;
KW central nervous system; cytostatic; apoptosis; ss.

XX Homo sapiens.

OS Location/Qualifiers

FT CDS 169..3246

FT /*tag= a

FT /product= "Human melanoma differentiation associated

FT (Mda)-5 protein"

FT misc_feature 3225..3228

FT /*tag= b

FT /note= "ATTTA motif"

FT misc_feature 3284..3287

FT /*tag= c

FT /note= "ATTTA motif"

FT polyA_signal 3343..3348

FT /*tag= d

XX WO200164707-A1.

PN 07-SEP-2001.

XX 28-FEB-2001; 2001WO-US006960.

XX 29-FEB-2000; 2000US-00515363.

XX (UYCO) UNIV COLUMBIA NEW YORK.

XX Fisher PB, Kang D, Gopalkrishnan RV;

XX

DR WPI; 2001-565494/63.
DR P-PSDB; AAE10155.
XX
PT Nucleic acid sequences encoding a Melanoma Differentiation Associated
PT Gene useful for cancer cell growth suppression, apoptosis and anti-viral
PT activity.
XX
PS Claim 1; Page 16-18; 152pp; English.
XX
CC The present invention relates to an isolated nucleic acid encoding a
CC melanoma differentiation associated gene-5 (Mda-5) polypeptide. Mda-5
CC contains a caspase recruitment domain (CARD) and a RNA helicase motif.
CC Mda-5 is a novel interferon (IFN) inducible gene with structural
CC similarities to RNA helicases and CARD motif containing proteins. Mda-5
CC is induced during terminal differentiation in human melanoma cells
CC treated with the combination of recombinant fibroblast IFN and the
CC antileukaemic compound mezerein (MEZ). Mda-5 is useful for identifying
CC compounds that may induce its expression. Mda-5 is useful for treating
CC cancer e.g., melanoma, neuroblastoma, astrocytoma, glioblastoma
CC multiforme, cervical cancer, breast cancer, colon cancer, prostate
CC cancer, osteosarcoma, chondrosarcoma and especially where the cancer is a
CC cancer of the central nervous system and apoptosis. The Mda-5 promoter
CC exhibits melanocyte tissue specificity and minimises systemic toxicity.
CC The present sequence is human Mda-5 cDNA
XX
SQ Sequence 3365 BP; 1138 A; 644 C; 753 G; 830 T; 0 U; 0 Other;
Query Match 100.0%; Score 3365; DB 4; Length 3365;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3365; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCGCGCCGGCCTGAGAGCCCTGTGGACAACTCGTCATTGTACGGCACAGAGCGGTAGAC 60
DB 1 GCGCGCCGGCCTGAGAGCCCTGTGGACAACTCGTCATTGTACGGCACAGAGCGGTAGAC 60
QY 61 CCTGCTTCTTAAGTGGGAGCGGACAGCGGCAGCGCACATTTCACTGTCCCGCAGACAA 120
DB 61 CCTGCTTCTTAAGTGGGAGCGGACAGCGGCAGCGCACATTTCACTGTCCCGCAGACAA 120
QY 121 CAGCACCATCTGCTTGGGAGAACCTCTCCCTTCTCTGAGAAAGAAAGATGTCGAATGGG 180
DB 121 CAGCACCATCTGCTTGGGAGAACCTCTCCCTTCTCTGAGAAAGAAAGATGTCGAATGGG 180
QY 181 TATTCACAGACGAGAAATTCGGCTATCTCATCTCGTCTTCCAGGGCCAGGGTGAATG 240
DB 181 TATTCACAGACGAGAAATTCGGCTATCTCATCTCGTCTTCCAGGGCCAGGGTGAATG 240
QY 241 TACATCCAGGTGGAGCCTGTGCTGGACTACCTGACCTTCTGCCTGCAGAGGTTGAAGGAG 300
DB 241 TACATCCAGGTGGAGCCTGTGCTGGACTACCTGACCTTCTGCCTGCAGAGGTTGAAGGAG 300
QY 301 CAGATTACAGGACAGTCGCCACCTCCGGGAACATGAGGAGTGAATGCTGCTGAGC 360
DB 301 CAGATTACAGGACAGTCGCCACCTCCGGGAACATGAGGAGTGAATGCTGCTGAGC 360
QY 361 ACCTTGGAGAGGGAGTCTGGCACCCTTGGTGGACTCGGGAATTCGTGAGGCCCCTCCGG 420
DB 361 ACCTTGGAGAGGGAGTCTGGCACCCTTGGTGGACTCGGGAATTCGTGAGGCCCCTCCGG 420
QY 421 AGAACCGGACGCCCTCTGGCCGCGCTACATGAACCTTGAGCTCACGGACTTGCCTCT 480
DB 421 AGAACCGGACGCCCTCTGGCCGCGCTACATGAACCTTGAGCTCACGGACTTGCCTCT 480
QY 481 CCATCGTTTGAGAACGCTCATGATGAATATCTCCAACCTGCTGAACCTCTTACGCCCACT 540
DB 481 CCATCGTTTGAGAACGCTCATGATGAATATCTCCAACCTGCTGAACCTCTTACGCCCACT 540
QY 541 CTGGTGGACAAAGCTTCTAGTTAGAGACCTCTTGGATAAGTGCATGGAGGAACTGTTG 600
DB 541 CTGGTGGACAAAGCTTCTAGTTAGAGACCTCTTGGATAAGTGCATGGAGGAACTGTTG 600
QY 601 ACAATTGAAGACAGAAACCGGATTGTGCTGCAGAAACAAATGGAAATGAATCAGGTGTA 660
DB 601 ACAATTGAAGACAGAAACCGGATTGTGCTGCAGAAACAAATGGAAATGAATCAGGTGTA 660

DB 601 ACAATTGAAGACAGAAACCGGATTGTGCTGCAGAAACAAATGGAAATGAATCAGGTGTA 660
QY 661 AGAGAGCTACTAAAAAGGATTGTGCAGAAAGAAACTGGTTCTCTGCAATTTCTGAATGTT 720
DB 661 AGAGAGCTACTAAAAAGGATTGTGCAGAAAGAAACTGGTTCTCTGCAATTTCTGAATGTT 720
QY 721 CTTTCGTAACACAGGAAACAAATGAACCTTGTCCAAGATTAAACAGGCTCTGATTTGCTCAGAA 780
DB 721 CTTTCGTAACACAGGAAACAAATGAACCTTGTCCAAGATTAAACAGGCTCTGATTTGCTCAGAA 780
QY 781 AGCAATGCAGAGATTGAGAAATTTATCACAAGTTTGTAGTTCAGAAATCAGACACAAGTTTGGCAGAA 840
DB 781 AGCAATGCAGAGATTGAGAAATTTATCACAAGTTTGTAGTTCAGAAATCAGACACAAGTTTGGCAGAA 840
QY 841 CTTTCAACCCACAGTTTCAGCCAAATCTGGAGAGGAGTCTGGGGCATGGAGATAACTCA 900
DB 841 CTTTCAACCCACAGTTTCAGCCAAATCTGGAGAGGAGTCTGGGGCATGGAGATAACTCA 900
QY 901 TCAGAATCATCTTTTGCAGATTCTTCTGTAGTTTCAGAAATCAGACACAAGTTTGGCAGAA 960
DB 901 TCAGAATCATCTTTTGCAGATTCTTCTGTAGTTTCAGAAATCAGACACAAGTTTGGCAGAA 960
QY 961 GGAAGTGTGAGTCTTAGATGAAAGTCTTGGACATAACAGCAACATGGGCGATGATTC 1020
DB 961 GGAAGTGTGAGTCTTAGATGAAAGTCTTGGACATAACAGCAACATGGGCGATGATTC 1020
QY 1021 GGCACCATGGGAAGTATTGAGATGAAGAGAATGTGGCAGCAAGAGCATCCCCGGAGCCA 1080
DB 1021 GGCACCATGGGAAGTATTGAGATGAAGAGAATGTGGCAGCAAGAGCATCCCCGGAGCCA 1080
QY 1081 GAACTCCAGCTCAGGCTTACCATAATGGAAGTTGCCAGCCAGCTTGGAGGGGAAGAT 1140
DB 1081 GAACTCCAGCTCAGGCTTACCATAATGGAAGTTGCCAGCCAGCTTGGAGGGGAAGAT 1140
QY 1141 ATCATCATCTGCTTCCCTACAGGGAGTGGAAAAACCCAGAGTGGCTGTTTACATTTGCCAAG 1200
DB 1141 ATCATCATCTGCTTCCCTACAGGGAGTGGAAAAACCCAGAGTGGCTGTTTACATTTGCCAAG 1200
QY 1201 GATCATTAGACAAGAAAGAAAGCATCTGAGCCTGGAAAGTTATAGTTCTTGTCAAT 1260
DB 1201 GATCATTAGACAAGAAAGAAAGCATCTGAGCCTGGAAAGTTATAGTTCTTGTCAAT 1260
QY 1261 AAGTACTGCTAGTTGAACAGCTCTTCCGCAAGGAGTTCCAAACCATTTTGAAGAAATGG 1320
DB 1261 AAGTACTGCTAGTTGAACAGCTCTTCCGCAAGGAGTTCCAAACCATTTTGAAGAAATGG 1320
QY 1321 TATCGTGTATTGGATTAAAGTGGTATACCCAACTGAAATATCATTTCCAGAAAGTTGTC 1380
DB 1321 TATCGTGTATTGGATTAAAGTGGTATACCCAACTGAAATATCATTTCCAGAAAGTTGTC 1380
QY 1381 AAGTCTGTGATATTATATCAGTACAGCTCAAACTCTTGAACCTCTTAAACTTG 1440
DB 1381 AAGTCTGTGATATTATATCAGTACAGCTCAAACTCTTGAACCTCTTAAACTTG 1440
QY 1441 GAAATGGAGAGAGATGCTGGTGTTCATTTGTCAGACTTTTCCCTCATTTATCATTTGATGAA 1500
DB 1441 GAAATGGAGAGAGATGCTGGTGTTCATTTGTCAGACTTTTCCCTCATTTATCATTTGATGAA 1500
QY 1501 TGTCTATCACCAACAAAGAGAGAGTGTATATAATACATCATGAGGCAATTTTGTGATGTCAG 1560
DB 1501 TGTCTATCACCAACAAAGAGAGAGTGTATATAATACATCATGAGGCAATTTTGTGATGTCAG 1560
QY 1561 AAGTTGAAAAACAATAGACTCAAGAAAGAAACAAACCAAGTATTCCCTTCTCAGATA 1620
DB 1561 AAGTTGAAAAACAATAGACTCAAGAAAGAAACAAACCAAGTATTCCCTTCTCAGATA 1620
QY 1621 CTGGGACTAACAGCTTACCTGGTGTGGAGGGGCCAGGAAGCAAGCCAAAGCTGAAGAA 1680
DB 1621 CTGGGACTAACAGCTTACCTGGTGTGGAGGGGCCAGGAAGCAAGCCAAAGCTGAAGAA 1680
QY 1681 CACATTTTAAACTATGTGCCAATCTTGTGATGCAATTTACTATTAAAACTGTTTAAAGAAAC 1740
DB 1681 CACATTTTAAACTATGTGCCAATCTTGTGATGCAATTTACTATTAAAACTGTTTAAAGAAAC 1740

QY	1741	CTTGATCAACTGAAAAACCAAATACAGGAGCCATGCAAGAAAGTTTGCCATTGCAGATGCA	1800
Db	1741		
QY	1801	CTTGATCAACTGAAAAACCAAATACAGGAGCCATGCAAGAAAGTTTGCCATTGCAGATGCA	1800
Db	1801		
QY	1801	ACCAGAGAAGATCCATTTAAAGAGAAAACCTTCTAGAAAATAATGACAAGGATTCAAACCTTAT	1860
Db	1801		
QY	1861	ACCAGAGAAGATCCATTTAAAGAGAAAACCTTCTAGAAAATAATGACAAGGATTCAAACCTTAT	1860
Db	1861		
QY	1861	TGTCAAATGAGTCCAAATGTGAGATTTTGGAACCTCAACCCCTATGAACAATGGGCCATTCAA	1920
Db	1861		
QY	1861	TGTCAAATGAGTCCAAATGTGAGATTTTGGAACCTCAACCCCTATGAACAATGGGCCATTCAA	1920
Db	1861		
QY	1921	ATGGAAAAAAGCTGCAAAAAAAGGAAATCGCAAAAGAACGTGTTTGTGCAGAACATTGTG	1980
Db	1921		
QY	1921	ATGGAAAAAAGCTGCAAAAAAAGGAAATCGCAAAAGAACGTGTTTGTGCAGAACATTGTG	1980
Db	1921		
QY	1981	AGGAAGTACAATGAGGCCCTACAAATTAATGACACAATTCGAATGATAGATGCGTATACT	2040
Db	1981		
QY	1981	AGGAAGTACAATGAGGCCCTACAAATTAATGACACAATTCGAATGATAGATGCGTATACT	2040
Db	1981		
QY	2041	CATCTTGAAACCTTCTATAATGAAGAGAAAAGATAAGAAGTTTGCAGTCATAGAAGATGAT	2100
Db	2041		
QY	2041	CATCTTGAAACCTTCTATAATGAAGAGAAAAGATAAGAAGTTTGCAGTCATAGAAGATGAT	2100
Db	2041		
QY	2101	AGTGATGAGGGTGGTGATGATGAGTATTGTGATGGTGATGAAGATGAGGATGATTTAAAG	2160
Db	2101		
QY	2101	AGTGATGAGGGTGGTGATGATGAGTATTGTGATGGTGATGAAGATGAGGATGATTTAAAG	2160
Db	2101		
QY	2161	AAACCTTTGAAACCTGGATGAAAACAGATAGATTTCTCATGACTTTATTTTTTGAAAAACAAT	2220
Db	2161		
QY	2161	AAACCTTTGAAACCTGGATGAAAACAGATAGATTTCTCATGACTTTATTTTTTGAAAAACAAT	2220
Db	2161		
QY	2221	AAAATGTTGAAAAGGCTGGCTGAAAACCCAGAAATATGAAAAATGAAAAGCTGACCAAATTA	2280
Db	2221		
QY	2221	AAAATGTTGAAAAGGCTGGCTGAAAACCCAGAAATATGAAAAATGAAAAGCTGACCAAATTA	2280
Db	2221		
QY	2281	AGAAATACCATTAATGGAGCAATATACTAGGACTGAGGAATCAGCACGAGGAATAATCTTT	2340
Db	2281		
QY	2281	AGAAATACCATTAATGGAGCAATATACTAGGACTGAGGAATCAGCACGAGGAATAATCTTT	2340
Db	2281		
QY	2341	ACAAAAACACGACAGAGTGCATATGCGCTTTCCAGTGGATTACTGAAAAATGAAAAATTT	2400
Db	2341		
QY	2341	ACAAAAACACGACAGAGTGCATATGCGCTTTCCAGTGGATTACTGAAAAATGAAAAATTT	2400
Db	2341		
QY	2401	GCTGAAGTAGGAGTCAAAGCCCCACCATCTGATTGGAGCTGGACACAGCAGTGAGTTCAAA	2460
Db	2401		
QY	2401	GCTGAAGTAGGAGTCAAAGCCCCACCATCTGATTGGAGCTGGACACAGCAGTGAGTTCAAA	2460
Db	2401		
QY	2461	CCCATGACACAGAAATGAAACAAAAAGAAGTCATTAGTAAAAATTCGCACCTGGAAAAATCAAT	2520
Db	2461		
QY	2461	CCCATGACACAGAAATGAAACAAAAAGAAGTCATTAGTAAAAATTCGCACCTGGAAAAATCAAT	2520
Db	2461		
QY	2521	CTGCTTATCGCTACACAGTGGCAGAGAAGGTCGTGATATTAAAGAATGTAACATTGTT	2580
Db	2521		
QY	2521	CTGCTTATCGCTACACAGTGGCAGAGAAGGTCGTGATATTAAAGAATGTAACATTGTT	2580
Db	2521		
QY	2581	ATCCGTTATGGTCTCGTCACCAATGAAATAGCCATGGTCCAGGCCCGTGGTCGAGCCAGA	2640
Db	2581		
QY	2581	ATCCGTTATGGTCTCGTCACCAATGAAATAGCCATGGTCCAGGCCCGTGGTCGAGCCAGA	2640
Db	2581		
QY	2641	GCTGATGAGAGCACCTACGTCCTGGTTGCTCACAGTGGTTCAGGAGTTATCGAACATGAG	2700
Db	2641		
QY	2641	GCTGATGAGAGCACCTACGTCCTGGTTGCTCACAGTGGTTCAGGAGTTATCGAACATGAG	2700
Db	2641		
QY	2701	ACAGTTAATGATTTCCGAGAGAAGATGATGTATAAAGCTATACATTGTGTTCAAAATATG	2760
Db	2701		
QY	2701	ACAGTTAATGATTTCCGAGAGAAGATGATGTATAAAGCTATACATTGTGTTCAAAATATG	2760
Db	2701		
QY	2761	AAACCAGAGGAGTATGCTCATTAAGATTTTGGAAATTAAGATGCAAGTATAATCGAAAAAG	2820
Db	2761		
QY	2761	AAACCAGAGGAGTATGCTCATTAAGATTTTGGAAATTAAGATGCAAGTATAATCGAAAAAG	2820
Db	2761		

QY	2821	AAAAATGAAACCAAGAGAGAAATATTGCCAAGCATTAACAAGAAATAACCCATCACTAACT	2880
DB	2821	AAAAATGAAACCAAGAGAGAAATATTGCCAAGCATTAACAAGAAATAACCCATCACTAACT	2880
QY	2881	TTCTTTTGCRAAAACTGCAGTGTCTAGCTGTTCTGGGAAGATATCCATGTAATTGAG	2940
DB	2881	TTCTTTTGCRAAAACTGCAGTGTCTAGCTGTTCTGGGAAGATATCCATGTAATTGAG	2940
QY	2941	AAAAATGCATCACGTCAATATGACCCCAAGAAATTCAGGAACTTTACATTTGTAAGAGAAAAC	3000
DB	2941	AAAAATGCATCACGTCAATATGACCCCAAGAAATTCAGGAACTTTACATTTGTAAGAGAAAAC	3000
QY	3001	AAAGCACTGCAAAAGAAGTGTGCCGACTATCAAAATAAATGGTGAAATCATCTGCANAATGT	3060
DB	3001	AAAGCACTGCAAAAGAAGTGTGCCGACTATCAAAATAAATGGTGAAATCATCTGCANAATGT	3060
QY	3061	GGCCAGGCTTGGGGAACAATGATGGTGCAAAAAGGCTTAGATTTGCCCTTGCTCAAAATA	3120
DB	3061	GGCCAGGCTTGGGGAACAATGATGGTGCAAAAAGGCTTAGATTTGCCCTTGCTCAAAATA	3120
QY	3121	AGGAATTTTGTAGTGGTTTTCAAAAATAAATTCACAAAGAAACAATACAAAAGTGGGTA	3180
DB	3121	AGGAATTTTGTAGTGGTTTTCAAAAATAAATTCACAAAGAAACAATACAAAAGTGGGTA	3180
QY	3181	GAATTACCTATCACATTTCCCAATCTTGACTATTTCAGAAATGCTGTTTATTAGTGTATGAG	3240
DB	3181	GAATTACCTATCACATTTCCCAATCTTGACTATTTCAGAAATGCTGTTTATTAGTGTATGAG	3240
QY	3241	GATTAGCACTTGATTTGAAGATTCTTTTAAAATACTATCAGTTAAACATTTAAATATGATTA	3300
DB	3241	GATTAGCACTTGATTTGAAGATTCTTTTAAAATACTATCAGTTAAACATTTAAATATGATTA	3300
QY	3301	TGATTAATGTATTCAATTATGCTACAGAACTGACATAAGAATCAATAAATGATTGTTTTTA	3360
DB	3301	TGATTAATGTATTCAATTATGCTACAGAACTGACATAAGAATCAATAAATGATTGTTTTTA	3360
QY	3361	CTCTG	3365
DB	3361	CTCTG	3365

RESULT 2

ADJ74904

ID ADJ74904 standard: DNA: 3380 BP.

AC ADJ74904;

XX

DT 20-MAY-2004 (first entry)

XX

DE Marker gene SEQ ID NO:156.

XX

KW bronchial asthma; chronic obstructive pulmonary disease;

KW respiratory epithelial cell; interlele

KW gene therapy:

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

OS Homo sapiens.

6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525
 526
 527
 528
 529

PN EP1394274-A2.

2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019 2020 2021 2022 2023 2024 2025 2026 2027 2028 2029 2030 2031 2032 2033 2034 2035 2036 2037 2038 2039 2040 2041 2042 2043 2044 2045 2046 2047 2048 2049 2050 2051 2052 2053 2054 2055 2056 2057 2058 2059 2060 2061 2062 2063 2064 2065 2066 2067 2068 2069 2070 2071 2072 2073 2074 2075 2076 2077 2078 2079 2080 2081 2082 2083 2084 2085 2086 2087 2088 2089 2090 2091 2092 2093 2094 2095 2096 2097 2098 2099 2100 2101 2102 2103 2104 2105 2106 2107 2108 2109 2110 2111 2112 2113 2114 2115 2116 2117 2118 2119 2120 2121 2122 2123 2124 2125 2126 2127 2128 2129 2130 2131 2132 2133 2134 2135 2136 2137 2138 2139 2140 2141 2142 2143 2144 2145 2146 2147 2148 2149 2150 2151 2152 2153 2154 2155 2156 2157 2158 2159 2160 2161 2162 2163 2164 2165 2166 2167 2168 2169 2170 2171 2172 2173 2174 2175 2176 2177 2178 2179 2180 2181 2182 2183 2184 2185 2186 2187 2188 2189 2190 2191 2192 2193 2194 2195 2196 2197 2198 2199 2200 2201 2202 2203 2204 2205 2206 2207 2208 2209 2210 2211 2212 2213 2214 2215 2216 2217 2218 2219 2220 2221 2222 2223 2224 2225 2226 2227 2228 2229 2230 2231 2232 2233 2234 2235 2236 2237 2238 2239 2240 2241 2242 2243 2244 2245 2246 2247 2248 2249 2250 2251 2252 2253 2254 2255 2256 2257 2258 2259 2260 2261 2262 2263 2264 2265 2266 2267 2268 2269 2270 2271 2272 2273 2274 2275 2276 2277 2278 2279 2280 2281 2282 2283 2284 2285 2286 2287 2288 2289 2290 2291 2292 2293 2294 2295 2296 2297 2298 2299 2300 2301 2302 2303 2304 2305 2306 2307 2308 2309 2310 2311 2312 2313 2314 2315 2316 2317 2318 2319 2320 2321 2322 2323 2324 2325 2326 2327 2328 2329 2330 2331 2332 2333 2334 2335 2336 2337 2338 2339 2340 2341 2342 2343 2344 2345 2346 2347 2348 2349 2350 2351 2352 2353 2354 2355 2356 2357 2358 2359 2360 2361 2362 2363 2364 2365 2366 2367 2368 2369 2370 2371 2372 2373 2374 2375 2376 2377 2378 2379 2380 2381 2382 2383 2384 2385 2386 2387 2388 2389 2390 2391 2392 2393 2394 2395 2396 2397 2398 2399 2400 2401 2402 2403 2404 2405 2406 2407 2408 2409 2410 2411 2412 2413 2414 2415 2416 2417 2418 2419 2420 2421 2422 2423 2424 2425 2426 2427 2428 2429 2430 2431 2432 2433 2434 2435 2436 2437 2438 2439 2440 2441 2442 2443 2444 2445 2446 2447 2448 2449 2450 2451 2452 2453 2454 2455 2456 2457 2458 2459 2460 2461 2462 2463 2464 2465 2466 2467 2468 2469 2470 2471 2472 2473 2474 2475 2476 2477 2478 2479 2480 2481 2482 2483 2484 2485 2486 2487 2488 2489 2490 2491 2492 2493 2494 2495 2496 2497 2498 2499 2500 2501 2502 2503 2504 2505 2506 2507 2508 2509 2510 2511 2512 2513 2514 2515 2516 2517 2518 2519 2520 2521 2522 2523 2524 2525 2526 2527 2528 2529 2530 2531 2532 2533 2534 2535 2536 2537 2538 2539 2540 2541 2542 2543 2544 2545 2546 2547 2548 2549 2550 2551 2552 2553 2554 2555 2556 2557 2558 2559 2560 2561 2562 2563 2564 2565 2566 2567 2568 2569 2570 2571 2572 2573 2574 2575 2576 2577 2578 2579 2580 2581 2582 2583 2584 2585 2586 2587 2588 2589 2590 2591 2592 2593 2594 2595 2596 2597 2598 2599 2600 2601 2602 2603 2604 2605 2606 2607 2608 2609 2610 2611 2612 2613 2614 2615 2616 2617 2618 2619 2620 2621 2622 2623 2624 2625 2626 2627 2628 2629 2630 2631 2632 2633 2634 2635 2636 2637 2638 2639 2640 2641 2642 2643 2644 2645 2646 2647 2648 2649 2650 2651 2652 2653 2654 2655 2656 2657 2658 2659 2660 2661 2662 2663 2664 2665 2666 2667 2668 2669 2670 2671 2672 2673 2674 2675 2676 2677 2678 2679 2680 2681 2682 2683 2684 2685 2686 2687 2688 2689 2690 2691 2692 2693 2694 2695 2696 2697 2698 2699 2700 2701 2702 2703 2704 2705 2706 2707 2708 2709 2710 2711 2712 2713 2714 2715 2716 2717 2718 2719 2720 2721 2722 2723 2724 2725 2726 2727 2728 2729 2730 2731 2732 2733 2734 2735 2736 2737 2738 2739 2740 2741 2742 2743 2744 2745 2746 2747 2748 2749 2750 2751 2752 2753 2754 2755 2756 2757 2758 2759 2760 2761 2762 2763 2764 2765 2766 2767 2768 2769 2770 2771 2772 2773 2774 2775 2776 2777 2778 2779 2780 2781 2782 2783 2784 2785 2786 2787 2788 2789 2790 2791 2792 2793 2794 2795 2796 2797 2798 2799 2800 2801 2802 2803 2804 2805 2806 2807 2808 2809 2810 2811 2812 2813 2814 2815 2816 2817 2

03-MAR-2004.

XX
DE 04-ATC-2003-04-AUG-2003;
PF
YY

06-AUG-2002:

FR 06-AUG-2002;
PR 20-MAR-2003:

comparing the expression level of a marker gene in a biological sample from a subject with the expression level of the gene in a sample from a healthy subject.

Claim 1: SEO ID NO 156: 241pp; English.

The present invention describes a method of testing for bronchial asthma or chronic obstructive pulmonary disease. The method comprises determining the expression level of a marker gene in a biological sample from a subject, comparing the expression level determined with the expression level of the marker gene in a biological sample from a healthy subject, and judging whether the subject has bronchial asthma or chronic obstructive pulmonary disease. The marker gene comprises: (a) a group of genes (S1) whose expression levels increase when respiratory epithelial cells are stimulated with interleukin-13; or (b) a group of genes (S2) whose expression levels decrease when respiratory epithelial cells are stimulated with interleukin-13. Also described: (1) a reagent (I) for testing for bronchial asthma or chronic obstructive pulmonary disease; (2) a kit for screening for a candidate compound for a therapeutic agent to treat bronchial asthma or chronic obstructive pulmonary disease; (3) an animal model for bronchial asthma or chronic obstructive pulmonary disease; (4) an inducer that induces bronchial asthma in a mouse; (5) a method for producing an animal model for bronchial asthma or chronic obstructive pulmonary disease; (6) a therapeutic agent for bronchial asthma or chronic obstructive pulmonary disease, comprising the compound, a marker gene or an antisense nucleic acid corresponding to a portion of the marker gene, a ribozyme, a polynucleotide that suppresses the expression of the gene through an RNAi effect or an antibody recognising a protein encoded by a marker gene; and (7) a DNA chip for testing for bronchial asthma or a chronic obstructive pulmonary disease, on which a probe has been immobilised to assay a marker gene. (I) has respiratory and antiasthmatic activities, and can be used in gene therapy. The method is useful for testing for or screening for a therapeutic agent for bronchial asthma or chronic obstructive pulmonary disease. The present invention is used in the exemplification of the present invention.

Sequence 3380 BP: 1153 A; 644 C; 753 G; 830 T; 0 U; 0 Other;

Query Match	100.0%;	Score 3365;	DB 12;	Length 3380;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 3365;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;

QY	1	GGCGCCGGCCTGAGAGCCCTGTGGACAAACCTCGTCATTTGT	CAGGCA	CAGAGCGGTAGAC	60
Db	1	GGCGCCGGCCTGAGAGCCCTGTGGACAAACCTCGTCATTTGT	CAGGCA	CAGAGCGGTAGAC	60
QY	61	CCTGCTTCTCTAAGTGGCAGCGGACAGCGGCACGACATTTCA	CTCTCCGC	CAGACAA	120
Db	61	CCTGCTTCTCTAAGTGGCAGCGGACAGCGGCACGACATTTCA	CTCTCCGC	CAGACAA	120
QY	121	CAGCACCATCTGCTTGGGAGAACCTCTCCCTTCTCTGAGAAAGAA	AGATGT	CGAATGGG	180
Db	121	CAGCACCATCTGCTTGGGAGAACCTCTCCCTTCTCTGAGAAAGAA	AGATGT	CGAATGGG	180
QY	181	TATTTCCACAGACGAGAATTTCCGCTATCTCATCTCGTGCTT	CAGGGCC	CAGGTTGAAAATG	240
Db	181	TATTTCCACAGACGAGAATTTCCGCTATCTCATCTCGTGCTT	CAGGGCC	CAGGTTGAAAATG	240
QY	241	TACATCCAGGTGGAGCCTGTGCTGGACTACCTGACCTTTTCT	GTGCCTG	CAGAGGTGAAGGAG	300
Db	241	TACATCCAGGTGGAGCCTGTGCTGGACTACCTGACCTTTTCT	GTGCCTG	CAGAGGTGAAGGAG	300
QY	301	CAGATTCAGAGGACAGTCGCCACCTCCGGGAACATGCAGGCAG	TGAACTG	CTGCTGAGC	360
Db	301	CAGATTCAGAGGACAGTCGCCACCTCCGGGAACATGCAGGCAG	TGAACTG	CTGCTGAGC	360
QY	361	ACCTTGGAGAAAGGAGTCTGGCACCTTGGTTGGACTCGGGAA	TTCTGTGG	AGGCCCTCCGG	420
Db	361	ACCTTGGAGAAAGGAGTCTGGCACCTTGGTTGGACTCGGGAA	TTCTGTGG	AGGCCCTCCGG	420
QY	421	AGAACCGGCAGCCCTCTGGCCGCCCGCTACATGAACCTGAGT	CACGGAC	TTTGCCCTCT	480
Db	421	AGAACCGGCAGCCCTCTGGCCGCCCGCTACATGAACCTGAGT	CACGGAC	TTTGCCCTCT	480

QY 1561 AAGTTGAAAAACAATAGACTCAAGAAAGAAAAACCAACCAGTAGTTCCTCCCTTCTCAGATA 1620
Db |||||||
QY 1561 AAGTTGAAAAACAATAGACTCAAGAAAGAAAAACCAACCAGTAGTTCCTCCCTTCTCAGATA 1620
Db |||||||
QY 1621 CTGGGACTAACAGCTTCAACCTGGTGTGGAGGGGCCACGAAGCAAGCCAAAGCTGAAGAA 1680
Db |||||||
QY 1621 CTGGGACTAACAGCTTCAACCTGGTGTGGAGGGGCCACGAAGCAAGCCAAAGCTGAAGAA 1680
Db |||||||
QY 1681 CACATTTTAAAACTATGTGCAATCTTGATGCAATTTACTATTAAAACTGTTAAAGAAAAAC 1740
Db |||||||
QY 1681 CACATTTTAAAACTATGTGCAATCTTGATGCAATTTACTATTAAAACTGTTAAAGAAAAAC 1740
Db |||||||
QY 1741 CTTGATCAAACTGAAAAACCAAAATACAGGAGCCATGCAAGAAGTTTGCCATTGCAGATGCA 1800
Db |||||||
QY 1741 CTTGATCAAACTGAAAAACCAAAATACAGGAGCCATGCAAGAAGTTTGCCATTGCAGATGCA 1800
Db |||||||
QY 1801 ACCAGAGAAGATCCATTTTAAAGAGAAAACCTTCTAGAAAATAATGACAAGGATTCAAACTTTAT 1860
Db |||||||
QY 1801 ACCAGAGAAGATCCATTTTAAAGAGAAAACCTTCTAGAAAATAATGACAAGGATTCAAACTTTAT 1860
Db |||||||
QY 1861 TGTCAATGAGTCCAAATGTCAGATTTTGGAACTCAACCCCTATGAACAATGGGCCATTCAA 1920
Db |||||||
QY 1861 TGTCAATGAGTCCAAATGTCAGATTTTGGAACTCAACCCCTATGAACAATGGGCCATTCAA 1920
Db |||||||
QY 1921 ATGGAAAAAAAGCTGCAAAAAAGGAAATCGCAAGAACGTTGTTGTCAGAAACATTTG 1980
Db |||||||
QY 1921 ATGGAAAAAAAGCTGCAAAAAAGGAAATCGCAAGAACGTTGTTGTCAGAAACATTTG 1980
Db |||||||
QY 1981 AGGAAGTACAATGAGGCCCTACAAATTAATGACACAATTCGAATGATAGATGCGTATACT 2040
Db |||||||
QY 1981 AGGAAGTACAATGAGGCCCTACAAATTAATGACACAATTCGAATGATAGATGCGTATACT 2040
Db |||||||
QY 2041 CATCTTGAAACTTTCTATAATGAAGAGAAAGATAAGAAAGTTTGCAGTCATAGAAGATGAT 2100
Db |||||||
QY 2041 CATCTTGAAACTTTCTATAATGAAGAGAAAGATAAGAAAGTTTGCAGTCATAGAAGATGAT 2100
Db |||||||
QY 2101 AGTGATGAGGGTGGTGATGAGTATTGTGATGGTGATGAAGATGAGGATGATTTAAAG 2160
Db |||||||
QY 2101 AGTGATGAGGGTGGTGATGAGTATTGTGATGGTGATGAAGATGAGGATGATTTAAAG 2160
Db |||||||
QY 2161 AAACCTTTGAAACTGGATGAAAACAGATAGATTTCTCATGACTTTATTTTGTGAAAAACAAT 2220
Db |||||||
QY 2161 AAACCTTTGAAACTGGATGAAAACAGATAGATTTCTCATGACTTTATTTTGTGAAAAACAAT 2220
Db |||||||
QY 2221 AAAATGTTGAAAAGGCTGGCTGAAAAACCCAGAAATATGAAAATGAAAAGCTGACCAAAATTA 2280
Db |||||||
QY 2221 AAAATGTTGAAAAGGCTGGCTGAAAAACCCAGAAATATGAAAATGAAAAGCTGACCAAAATTA 2280
Db |||||||
QY 2281 AGAAAATACCATAATGGAGCAATATACTAGGACTGAGGAATCAGCAGAGGAATAATCTTT 2340
Db |||||||
QY 2281 AGAAAATACCATAATGGAGCAATATACTAGGACTGAGGAATCAGCAGAGGAATAATCTTT 2340
Db |||||||
QY 2341 ACAAAAAACGACAGAGTGCATATGCGCTTTCCAGTGGATTACTGAAAAATGAAAAATTT 2400
Db |||||||
QY 2341 ACAAAAAACGACAGAGTGCATATGCGCTTTCCAGTGGATTACTGAAAAATGAAAAATTT 2400
Db |||||||
QY 2401 GCTGAAGTAGGAGTCAAAGCCCCACCATCTGATTGGAGCTGGACACAGCAGTGAGTTCAA 2460
Db |||||||
QY 2401 GCTGAAGTAGGAGTCAAAGCCCCACCATCTGATTGGAGCTGGACACAGCAGTGAGTTCAA 2460
Db |||||||
QY 2461 CCCATGACACAGAAATGAACAAAAAGAAAGTCAATAGTAAATTTTCGCACTGAAAAATCAAT 2520
Db |||||||
QY 2461 CCCATGACACAGAAATGAACAAAAAGAAAGTCAATAGTAAATTTTCGCACTGAAAAATCAAT 2520
Db |||||||
QY 2521 CTGCTTATCGCTACCCAGTGGCAGAAAGCTCTGGATATTAAAGAATGTAACATTGTT 2580
Db |||||||
QY 2521 CTGCTTATCGCTACCCAGTGGCAGAAAGCTCTGGATATTAAAGAATGTAACATTGTT 2580
Db |||||||
QY 2581 ATCCGTTATGGTCTCGTCACCAATGAAATAGCCATGGTCCAGGCCCGTGGTCGAGCCAGA 2640
Db |||||||
QY 2581 ATCCGTTATGGTCTCGTCACCAATGAAATAGCCATGGTCCAGGCCCGTGGTCGAGCCAGA 2640
Db |||||||
QY 2641 GCTGATGAGAGCACCTACGTCCTGTTGCTCAGAGTGGTTTCAGGAGTTATCGAACATGAG 2700

Db |||||||
2641 GCTGATGAGAGCACCTACGTCCTGGTTGCTCACAGTGGTTCAGGAGTTATCGAACATGAG 2700
QY |||||||
2701 ACAGTTAATGATTTCCGAGAGAAAGATGATGTATAAAGCTATACATTGTGTTCAAAATATG 2760
Db |||||||
2701 ACAGTTAATGATTTCCGAGAGAAAGATGATGTATAAAGCTATACATTGTGTTCAAAATATG 2760
QY |||||||
2761 AAACCAGAGGAGTATGCTCATAGATTTTGGAAATTACAGATGCAAGATATAATGGAAAAAG 2820
Db |||||||
2761 AAACCAGAGGAGTATGCTCATAGATTTTGGAAATTACAGATGCAAGATATAATGGAAAAAG 2820
QY |||||||
2821 AAAATGAAAAACCAAGAGAAAAATATTGCCAAGCATTACAAGAAATAACCCCATCACTAATACT 2880
Db |||||||
2821 AAAATGAAAAACCAAGAGAAAAATATTGCCAAGCATTACAAGAAATAACCCCATCACTAATACT 2880
QY |||||||
2881 TTCTTTTGCAAAAACTGCAGTGTGCTAGCCTGTTCTGGGGAAGATATCCATGTAATTGAG 2940
Db |||||||
2881 TTCTTTTGCAAAAACTGCAGTGTGCTAGCCTGTTCTGGGGAAGATATCCATGTAATTGAG 2940
QY |||||||
2941 AAAATGCATCACGTCAATATGACCCCAAGAACTTCAAGGAACTTTACATTGTAAGAGAAAAAC 3000
Db |||||||
2941 AAAATGCATCACGTCAATATGACCCCAAGAACTTCAAGGAACTTTACATTGTAAGAGAAAAAC 3000
QY |||||||
3001 AAAGCACTGCAAAAGAAAGTGTGCCGACTATCAAAATAAATGTTGTAATCATCTGCAAAATGT 3060
Db |||||||
3001 AAAGCACTGCAAAAGAAAGTGTGCCGACTATCAAAATAAATGTTGTAATCATCTGCAAAATGT 3060
QY |||||||
3061 GGCCAGGCTTGGGGAACAATGATGGTGCAAAAAGGCTTAGATTGCTCTCAAAAATA 3120
Db |||||||
3061 GGCCAGGCTTGGGGAACAATGATGGTGCAAAAAGGCTTAGATTGCTCTCAAAAATA 3120
QY |||||||
3121 AGGAATTTTGTAGTGGTTTTTCAAAAAATAAATCAACAAAGAAACAATACAAAAAAGTGGGTA 3180
Db |||||||
3121 AGGAATTTTGTAGTGGTTTTTCAAAAAATAAATCAACAAAGAAACAATACAAAAAAGTGGGTA 3180
QY |||||||
3181 GAATTACCTATCACATTTCCCAATCTTGACTATTTCAGAAATGCTGTTTATTATTAGTGATGAG 3240
Db |||||||
3181 GAATTACCTATCACATTTCCCAATCTTGACTATTTCAGAAATGCTGTTTATTATTAGTGATGAG 3240
QY |||||||
3241 GATTAGCACTTGATTGAAGATTCTTTTAAAAATACTATCAGTTAAACATTTAATATGATTA 3300
Db |||||||
3241 GATTAGCACTTGATTGAAGATTCTTTTAAAAATACTATCAGTTAAACATTTAATATGATTA 3300
QY |||||||
3301 TGATTAATGTATTCATTATGCTACAGAACTGACATAAAGATCAATAAAAAATGATTGTTTAA 3360
Db |||||||
3301 TGATTAATGTATTCATTATGCTACAGAACTGACATAAAGATCAATAAAAAATGATTGTTTAA 3360
QY |||||||
3361 CTCTG 3365
Db |||||||
3361 CTCTG 3365

RESULT 3

ADN04879
ID ADN04879 standard; cDNA; 3379 BP.
XX
AC ADN04879;
XX
DT 01-JUL-2004 (first entry)
XX
DE Antipsoriatic cDNA sequence #654.
XX
KW ds; gene; antipsoriatic; gene therapy; psoriasis; diagnosis.
XX
OS Homo sapiens.
XX
PN WO2004028479-A2.
-XX
PD 08-APR-2004.
XX
PF 25-SEP-2003; 2003WO-US030907.
XX
PR 25-SEP-2002; 2002US-0414006P.

QY 1801 ACCAGAGAAGATCCATTATAAGAGAGAACTTCTAGAAAATAATGACAAGGATTCAAACCTTAT 1860
Db 1800 ACCAGAGAAGATCCATTATAAGAGAGAACTTCTAGAAAATAATGACAAGGATTCAAACCTTAT 1859
QY 1861 TGTCAAATGAGTCCCAATGTCCAGATTTTGGAACTCAACCCCTATGAACAATGGGCCATTCAA 1920
Db 1860 TGTCAAATGAGTCCCAATGTCCAGATTTTGGAACTCAACCCCTATGAACAATGGGCCATTCAA 1919
QY 1921 ATGGAAAAAAGCTGCAAAAAAAGGAAATCGCAAAAGAACGTGTTGTGCAGAACATTTG 1980
Db 1920 ATGGAAAAAAGCTGCAAAAAAAGGAAATCGCAAAAGAACGTGTTGTGCAGAACATTTG 1979
QY 1981 AGGAAGTACAATGAGGCCCTACAAATTAATGACACAATTCGAATGATAGATGCGTATACT 2040
Db 1980 AGGAAGTACAATGAGGCCCTACAAATTAATGACACAATTCGAATGATAGATGCGTATACT 2039
QY 2041 CATCTTGAACCTTTCTATAATGAAGAGAAAGATAAGAAAGTTTGCAGTCATAGAAGATGAT 2100
Db 2040 CATCTTGAACCTTTCTATAATGAAGAGAAAGATAAGAAAGTTTGCAGTCATAGAAGATGAT 2099
QY 2101 AGTGATGAGGGTGGTGATGATGAGTATTGTGATGGTGATGAAGATGAGGATGATTTAAAG 2160
Db 2100 AGTGATGAGGGTGGTGATGATGAGTATTGTGATGGTGATGAAGATGAGGATGATTTAAAG 2159
QY 2161 AAACCTTTGAAACTGGATGAAACAGATAGATTTCTCATGACTTTATTTTGGAAAAACAAT 2220
Db 2160 AAACCTTTGAAACTGGATGAAACAGATAGATTTCTCATGACTTTATTTTGGAAAAACAAT 2219
QY 2221 AAAATGTTGAAAAGGCTGGCTGAAAACCCAGAAATATGAAAATGAAAAGCTGACCAAAATTA 2280
Db 2220 AAAATGTTGAAAAGGCTGGCTGAAAACCCAGAAATATGAAAATGAAAAGCTGACCAAAATTA 2279
QY 2281 AGAAATACCATATAATGGAGCAATATATACTAGGACTGAGGAATCAGCAGGGAATAATCTTT 2340
Db 2280 AGAAATACCATATAATGGAGCAATATATACTAGGACTGAGGAATCAGCAGGGAATAATCTTT 2339
QY 2341 ACAAAAAACGACAGAGTGCCATATCGCTTTCCAGTGGATTACTGAAAAATGAAAAATTT 2400
Db 2340 ACAAAAAACGACAGAGTGCCATATCGCTTTCCAGTGGATTACTGAAAAATGAAAAATTT 2399
QY 2401 GCTGAAGTAGGAGTCAAAAGCCCAACCATCTGATTGGAGCTGGACACAGCAGTGAGTTCAA 2460
Db 2400 GCTGAAGTAGGAGTCAAAAGCCCAACCATCTGATTGGAGCTGGACACAGCAGTGAGTTCAA 2459
QY 2461 CCCATGACACAGAAATGAACAAAAAGAACGATAGTAAATTTCCGACTGGAAAAATCAAT 2520
Db 2460 CCCATGACACAGAAATGAACAAAAAGAACGATAGTAAATTTCCGACTGGAAAAATCAAT 2519
QY 2521 CTGCTTATCGTACCAAGTGGCAGAGAAAGTCTGGATATTAAAGAAATGTAACATTTGT 2580
Db 2520 CTGCTTATCGTACCAAGTGGCAGAGAAAGTCTGGATATTAAAGAAATGTAACATTTGT 2579
QY 2581 ATCCGTTATGGTCTCGTCACCAATGAAATAGCCATGGTCCAGGCCGTGGTCGAGCCAGA 2640
Db 2580 ATCCGTTATGGTCTCGTCACCAATGAAATAGCCATGGTCCAGGCCGTGGTCGAGCCAGA 2639
QY 2641 GCTGATGAGAGCACCTACGTCCTGGTTGCTCACAGTGGTTCAGGAGTTATCGAACATGAG 2700
Db 2640 GCTGATGAGAGCACCTACGTCCTGGTTGCTCACAGTGGTTCAGGAGTTATCGAACATGAG 2699
QY 2701 ACAGTTAATGATTTCCGAGAGAAAGATGATGATATAAGCTATACATTTGTTCAAAAATATG 2760
Db 2700 ACAGTTAATGATTTCCGAGAGAAAGATGATGATATAAGCTATACATTTGTTCAAAAATATG 2759
QY 2761 AAACCAGAGGAGTATGCTCATAAGATTTTGGAAATTAAGATGCAAGTATAATGGAAAG 2820
Db 2760 AAACCAGAGGAGTATGCTCATAAGATTTTGGAAATTAAGATGCAAGTATAATGGAAAG 2819
QY 2821 AAAATGAAAAACCAAGAGAAATATTGCCAAGCATTACAAGAATAACCCATCACTAATAACT 2880
Db 2820 AAAATGAAAAACCAAGAGAAATATTGCCAAGCATTACAAGAATAACCCATCACTAATAACT 2879

QY 2881 TTCTTTTGCAAAAACTGCAGTGTGCTAGCCTGTTCTGGGGAAGATATCCATGTAAATTGAG 2940
Db 2880 TTCTTTTGCAAAAACTGCAGTGTGCTAGCCTGTTCTGGGGAAGATATCCATGTAAATTGAG 2939
QY 2941 AAAATGCATCAGTCAATATGACCCAGAAATTCAGGAACCTTTACATTTGAAGAGAAAAAC 3000
Db 2940 AAAATGCATCAGTCAATATGACCCAGAAATTCAGGAACCTTTACATTTGAAGAGAAAAAC 2999
QY 3001 AAAGCACTGCAAAAAGAGTGTGCCGACTATCAAAATAAATGGTGAAATCATCTGCAAAATGT 3060
Db 3000 AAAGCACTGCAAAAAGAGTGTGCCGACTATCAAAATAAATGGTGAAATCATCTGCAAAATGT 3059
QY 3061 GGCCAGGCTTGGGAAACAATGATGCTGCACAAAGGCTTAGATTTGCCCTTGTCTCAAAATA 3120
Db 3060 GGCCAGGCTTGGGAAACAATGATGCTGCACAAAGGCTTAGATTTGCCCTTGTCTCAAAATA 3119
QY 3121 AGGAATTTTGTAGTGGTTTTTCAAAAATAAATTCACAAAGAAACAATACAAAAAGTGGGTA 3180
Db 3120 AGGAATTTTGTAGTGGTTTTTCAAAAATAAATTCACAAAGAAACAATACAAAAAGTGGGTA 3179
QY 3181 GAATTACCTATCACATTTCCCAATCTTGACTATTTCAGAAATGCTGTTTATTAGTGATGAG 3240
Db 3180 GAATTACCTATCACATTTCCCAATCTTGACTATTTCAGAAATGCTGTTTATTAGTGATGAG 3239
QY 3241 GATTAGCACTTGATTGAAGATTCTTTTAAAAATACTATCAGTTAAACATTTAATATGATTA 3300
Db 3240 GATTAGCACTTGATTGAAGATTCTTTTAAAAATACTATCAGTTAAACATTTAATATGATTA 3299
QY 3301 TGATTAATGTATTTCATTATGCTACAGAACTGACATAAGAAATCAATAAAATGATTGTTTTA 3360
Db 3300 TGATTAATGTATTTCATTATGCTACAGAACTGACATAAGAAATCAATAAAATGATTGTTTTA 3359
QY 3361 CTCTG 3365
Db 3360 CTCTG 3364

RESULT 4
ADQ22237
ID ADQ22237 standard; DNA; 3668 BP.
XX
AC ADQ22237;
XX
DT 26-AUG-2004 (first entry)
XX
DE Human soft tissue sarcoma-upregulated DNA - SEQ ID 5057.
XX
KW soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human;
KW ds.
OS Homo sapiens.
XX
PN WO2004048938-A2.
PD 10-JUN-2004.
XX
PF 26-NOV-2003; 2003WO-US038193.
XX
PR 26-NOV-2002; 2002US-0429739P.
XX
PA (PROT-) PROTEIN DESIGN LABS INC.
XX
PI Aziz N, Ginsburg WM, Zlotnik A;
XX
DR WPI; 2004-441208/41.
XX
PT Early detection of soft tissue sarcoma comprises determining expression
PT of a gene in a first soft tissue sample and a normal soft tissue sample
PT and comparing the gene expression, also useful in treating soft tissue
PT sarcoma.
PS Example 2; SEQ ID NO 5057; 210pp; English.
XX

CC The invention relates to a novel method for detecting soft tissue sarcoma
CC which comprises obtaining a first soft tissue sample from an individual
CC and a normal soft tissue sample from the same or different individual,
CC determining the expression of a gene in both samples and comparing the
CC expression of the gene in both soft tissue samples, where a higher level
CC of protein expression in the first soft tissue sample indicates the
CC presence of soft tissue sarcoma. The method of the invention has
CC cytotatic applications and may be useful for detecting soft tissue
CC sarcoma, possibly via gene therapy or vaccine production. The nucleic
CC acid sequences may be useful in diagnostic and screening applications.
CC The current sequence is that of a human soft tissue sarcoma-upregulated
CC DNA of the invention. The current sequence is not shown within the
CC specification per se but was submitted in CD format by the inventor.

XX
SQ Sequence 3668 BP; 1215 A; 730 C; 811 G; 872 T; 0 U; 40 Other;

Query Match 99.2%; Score 3339.2; DB 12; Length 3668;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 3341; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

Qy 1 GCGGCGCGCCTGAGAGCCCTGTGGACAACTCGTCATTGTCAGGCACAGAGCGGTAGAC 60
Dy
Qy 61 CCTGCTTCTTAAGTGGGACGGGACAGCGGCACGCACATTTACCTGTGCCGACAGCAA 120
Dy 61 CCTGCTTCTTAAGTGGGACGGGACAGCGGCACGCACATTTACCTGTGCCGACAGCAA 120
Qy 121 CAGCACCATCTGCTTGGGAGAACCTCTCCCTTCTCTGAGAAAGAAAGATGTCGAATGGG 180
Dy 121 CAGCACCATCTGCTTGGGAGAACCTCTCCCTTCTCTGAGAAAGAAAGATGTCGAATGGG 180
Qy 181 TATTCACAGACGAGAAATTCGGCTATCTCATCTCGTCTTACGGCCAGGGTGAAATG 240
Dy 181 TATTCACAGACGAGAAATTCGGCTATCTCATCTCGTCTTACGGCCAGGGTGAAATG 240
Qy 241 TACATCCAGGTGGAGCCTGTGCTGGACTACCTGACCTTCTGCTGCAGAGGTGAAGAG 300
Dy 241 TACATCCAGGTGGAGCCTGTGCTGGACTACCTGACCTTCTGCTGCAGAGGTGAAGAG 300
Qy 301 CAGATTACAGAGGACAGTCCGACCTCCGGGAAATGCAGGACAGTTGAACCTGCTGAGC 360
Dy 301 CAGATTACAGAGGACAGTCCGACCTCCGGGAAATGCAGGACAGTTGAACCTGCTGAGC 360
Qy 361 ACCTTGGAGAAGGAGTCTGGACCTTGGTGGACTCGGAAATTCGTGGAGGCCCTCCGG 420
Dy 361 ACCTTGGAGAAGGAGTCTGGACCTTGGTGGACTCGGAAATTCGTGGAGGCCCTCCGG 420
Qy 421 AGAACCGGACGCTCTGCGCGCGGCTACATGAACCTGAGCTCAGGACTTGCCTCT 480
Dy 421 AGAACCGGACGCTCTGCGCGCGGCTACATGAACCTGAGCTCAGGACTTGCCTCT 480
Qy 481 CCATCGTTTGAGAACGCTCATGATGAATATCTCCAACCTGTAACCTCCTTCAGCCCACT 540
Dy 481 CCATCGTTTGAGAACGCTCATGATGAATATCTCCAACCTGTAACCTCCTTCAGCCCACT 540
Qy 541 CTGGTGGACAAAGCTTCTAGTTAGAGACGCTTGGATAAGTGCATGGAGGAGAACTGTTG 600
Dy 541 CTGGTGGACAAAGCTTCTAGTTAGAGACGCTTGGATAAGTGCATGGAGGAGAACTGTTG 600
Qy 601 ACAATTGAAGACAGAAACCGGATTTGCTGCTGCAGAAACAAATGGAATGAATCAGGTGTA 660
Dy 601 ACAATTGAAGACAGAAACCGGATTTGCTGCTGCAGAAACAAATGGAATGAATCAGGTGTA 660
Qy 661 AGAGAGCTACTAAAAGGATTTGTGCAAGAAAGAAACTGGTTCTCTGCAATTTCTGAATGTT 720
Dy 661 AGAGAGCTACTAAAAGGATTTGTGCAAGAAAGAAACTGGTTCTCTGCAATTTCTGAATGTT 720
Qy 721 CTTCTGTAACACAGGAAACAAATGAACCTTGTCCAGAGTTAAGAGGCTCTGATTTGCTCAGAA 780
Dy 721 NNNNNNNNNNNNNNNNNATGAACCTTGTCCAAAGAGTTAAGAGGCTCTGATTTGCTCAGAA 780
Qy 781 AGCAATGCAGAGATTGAGAAATTTATCACAAGTTGATGGTCTCCTCAAGTGAAGAGCAACTT 840

Db 781 AGCAATGCAGAGATTGAGAAATTTATCACAAGTTGATGGTCTCCTCAAGTGAAGAGCAACTT 840
Qy 841 CTTTCAACCACAGTTTCAGCCAAATCTGGAGAGAGAGGTCTGGGGCATGGAGAAATAACTCA 900
Db 841 CTTTCAACCACAGTTTCAGCCAAATCTGGAGAGAGAGGTCTGGGGCATGGAGAAATAACTCA 900
Qy 901 TCAGAAATCATCTTTTTCAGAGATCTTCTGTAGTTTCAGAAATCAGACACAAAGTTTGGCAGAA 960
Db 901 TCAGAAATCATCTTTTTCAGAGATCTTCTGTAGTTTCAGAAATCAGACACAAAGTTTGGCAGAA 960
Qy 961 GGAAGTGTGAGCTGCTTAGATGAAAGTCTTGGACATAACAGCAACATGGGAGTGAATCA 1020
Db 961 GGAAGTGTGAGCTGCTTAGATGAAAGTCTTGGACATAACAGCAACATGGGAGTGAATCA 1020
Qy 1021 GGCACCATGGGAAGTGAATTCAGATGAAGAGAAATGTGGCAGCAAGAGCATCCCGGAGCCA 1080
Db 1021 GGCACCATGGGAAGTGAATTCAGATGAAGAGAAATGTGGCAGCAAGAGCATCCCGGAGCCA 1080
Qy 1081 GAACTCCAGCTCAGGCCCTTACCAAATGGAAGTTGCCAGCCAGCCCTTGGAGGGAAGAAT 1140
Db 1081 GAACTCCAGCTCAGGCCCTTACCAAATGGAAGTTGCCAGCCAGCCCTTGGAGGGAAGAAT 1140
Qy 1141 ATCATCATCTGCTCCCTACAGGGAGTGGAAAAACCAAGAGTGGCTGTTTACATGTCCTCAAG 1200
Db 1141 ATCATCATCTGCTCCCTACAGGGAGTGGAAAAACCAAGAGTGGCTGTTTACATGTCCTCAAG 1200
Qy 1201 GATCACTTAGACAAGAAGAAAAAGCATCTGAGCCTGGAAAAAGTTATAGTTCTTGTCAAT 1260
Db 1201 GATCACTTAGACAAGAAGAAAAAGCATCTGAGCCTGGAAAAAGTTATAGTTCTTGTCAAT 1260
Qy 1261 AAGTACTGCTAGTTGAACAGCTCTTCCGCAAGGAGTTCCAAACCATTTTGAAGAAATGG 1320
Db 1261 AAGTACTGCTAGTTGAACAGCTCTTCCGCAAGGAGTTCCAAACCATTTTGAAGAAATGG 1320
Qy 1321 TATCGTGTATTGGATTAAAGTGGTATACCCAACTGAAAAATATCATTTCCAGAAAGTTGTC 1380
Db 1321 TATCGTGTATTGGATTAAAGTGGTATACCCAACTGAAAAATATCATTTCCAGAAAGTTGTC 1380
Qy 1381 AAGTCTGTGATATTATATCAGTACAGCTCAAAATCCTTGAANAACCTCCTCTTAAACTTG 1440
Db 1381 AAGTCTGTGATATTATATCAGTACAGCTCAAAATCCTTGAANAACCTCCTCTTAAACTTG 1440
Qy 1441 GAAATGGAGAAGATGCTGGTGTTCATTTGTTCAGACTTTTCCCTCATTTATCATTTGATGAA 1500
Db 1441 GAAATGGAGAAGATGCTGGTGTTCATTTGTTCAGACTTTTCCCTCATTTATCATTTGATGAA 1500
Qy 1501 TGTCATCACACCAACAAAGAACGAGTGTATATAACATCATGAGGCAATTTTGTATGTCAG 1560
Db 1501 TGTCATCACACCAACAAAGAACGAGTGTATATAACATCATGAGGCAATTTTGTATGTCAG 1560
Qy 1561 AAGTTGAAAAACAATAGACTCAAGAAAGAAACCAACCAAGTATTCCTCCTCAGATA 1620
Db 1561 AAGTTGAAAAACAATAGACTCAAGAAAGAAACCAACCAAGTATTCCTCCTCAGATA 1620
Qy 1621 CTGGGACTAACAGCTTACCTGCTGTTGGAGGGGCCACGAAAGCTGAAGAA 1680
Db 1621 CTGGGACTAACAGCTTACCTGCTGTTGGAGGGGCCACGAAAGCTGAAGAA 1680
Qy 1681 CACATTTTAAAAACTATGTGCCAATCTTGTATGATGCAATTTTAAAGAAAAAC 1740
Db 1681 CACATTTTAAAAACTATGTGCCAATCTTGTATGATGCAATTTTAAAGAAAAAC 1740
Qy 1741 CTTGATCAACTGAAAAACCAAAATACAGGAGCCATGCAAGAAAGTTTGGCATTGTCAGATGCA 1800
Db 1741 CTTGATCAACTGAAAAACCAAAATACAGGAGCCATGCAAGAAAGTTTGGCATTGTCAGATGCA 1800
Qy 1801 ACCAGAGAAGATCCATTTAAAGAGAAAACTTCTAGAAATTAATGACAAGGATTCAAACTTAT 1860
Db 1801 ACCAGAGAAGATCCATTTAAAGAGAAAACTTCTAGAAATTAATGACAAGGATTCAAACTTAT 1860
Qy 1861 TGTCAAATGAGTCCAAATGTGAGATTTTGGAACTCAACCTCATGAACAATGGGCCATTTCAA 1920

Db 1861 TGTCAAATGAGTCCCAATGTCCAGATTTTGGRACTCAACCCCTATGAACAATGGGCCATTCAA 1920

QY 1921 ATGGAAAAAAGCTGCAAAAAAAGGAATCGCAAAAGAACGTGTTGTGCGAGAACATTG 1980

Db 1921 ATGGAAAAAAGCTGCAAAAGAGGAATCGCAAAAGAACGTGTTGTGCGAGAACATTG 1980

QY 1981 AGGAAGTACAATGAGGCCCTACAAATTAATGACACAATTCGAATGATAGATCGGTATACT 2040

Db 1981 AGGAAGTACAATGAGGCCCTACAAATTAATGACACAATTCGAATGATAGATCGGTATACT 2040

QY 2041 CATCTTGAAACTTTCTATAATGAAGAGAAAGATAAGAGTTTGCAGTCATAGAAGATGAT 2100

Db 2041 CATCTTGAAACTTTCTATAATGAAGAGAAAGATAAGAGTTTGCAGTCATAGAAGATGAT 2100

QY 2101 AGTGATGAGGGTGGTGATGATGAGTATTGTGATGGTGATGAAGATGAGGATGATTTAAAG 2160

Db 2101 AGTGATGAGGGTGGTGATGATGAGTATTGTGATGGTGATGAAGATGAGGATGATTTAAAG 2160

QY 2161 AAACCTTTGAAACTGGATGAAAACAGATAGATTCTCATGACTTTATTTTGGAAAAACAAT 2220

Db 2161 AAACCTTTGAAACTGGATGAAAACAGATAGATTCTCATGACTTTATTTTGGAAAAACAAT 2220

QY 2221 AAAATGTTGAAAAGCGTGGCTGAAAACCCAGAAATATGAAAATGAAAAGCTGACCAAAATTA 2280

Db 2221 AAAATGTTGAAAAGCGTGGCTGAAAACCCAGAAATATGAAAATGAAAAGCTGACCAAAATTA 2280

QY 2281 AGAAATACCATAATGGAGCAATATACTAGACTGAGGAATCAGCACGAGGAATAATCTTT 2340

Db 2281 AGAAATACCATAATGGAGCAATATACTAGACTGAGGAATCAGCACGAGGAATAATCTTT 2340

QY 2341 ACAAAAAACGACAGAGTGACATATGCGCTTTCCAGTGGATTACTGAAAAATGAAAAATTT 2400

Db 2341 ACAAAAAACGACAGAGTGACATATGCGCTTTCCAGTGGATTACTGAAAAATGAAAAATTT 2400

QY 2401 GCTGAAGTAGGAGTCAAAGCCCCACCATCTGATTGGAGCTGGACACAGCAGTGAGTTCAA 2460

Db 2401 GCTGAAGTAGGAGTCAAAGCCCCACCATCTGATTGGAGCTGGACACAGCAGTGAGTTCAA 2460

QY 2461 CCCATGACACAGAAATGAACAAAAAGAGTCAATTAGTAAATTTCCGACTGGAAAAATCAAT 2520

Db 2461 CCCATGACACAGAAATGAACAAAAAGAGTCAATTAGTAAATTTCCGACTGGAAAAATCAAT 2520

QY 2521 CTGCTTATCGCTACACAGTGGCAGAGAGGTCTGGATATTAAGAATGTAACATTGTT 2580

Db 2521 CTGCTTATCGCTACACAGTGGCAGAGAGGTCTGGATATTAAGAATGTAACATTGTT 2580

QY 2581 ATCCGTTATGGTCTCGTCACCAATGAAATAGCCATGGTCCAGGCCGTGGTCGAGCCAGA 2640

Db 2581 ATCCGTTATGGTCTCGTCACCAATGAAATAGCCATGGTCCAGGCCGTGGTCGAGCCAGA 2640

QY 2641 GCTGATGAGAGCACCTACGTCCTGTTGCTCACAGTGGTTCAGGAGTTATCGAACATGAG 2700

Db 2641 GCTGATGAGAGCACCTACGTCCTGTTGCTCACAGTGGTTCAGGAGTTATCGAACATGAG 2700

QY 2701 ACAGTTAATGATTTCCGAGAGAAAGATGATGATAAAGCTATACATTGTGTTCAAAATATG 2760

Db 2701 ACAGTTAATGATTTCCGAGAGAAAGATGATGATAAAGCTATACATTGTGTTCAAAATATG 2760

QY 2761 AAACCCAGAGGAGTATGCTCATAGATTTTGGAAATTACAGATGCAAAAGTATAATGGAAAAG 2820

Db 2761 AAACCCAGAGGAGTATGCTCATAGATTTTGGAAATTACAGATGCAAAAGTATAATGGAAAAG 2820

QY 2821 AAAATGAAAACCCAGAGAAATATTGCCAAGCATTACAAGAATAACCCCATCACTAATACT 2880

Db 2821 AAAATGAAAACCCAGAGAAATATTGCCAAGCATTACAAGAATAACCCCATCACTAATACT 2880

QY 2881 TTCTTTTGCAAAAACCTGAGTGTGCTAGCTGTTCTGGGGAAGATATCCATGTAATTGAG 2940

Db 2881 TTCTTTTGCAAAAACCTGAGTGTGCTAGCTGTTCTGGGGAAGATATCCATGTAATTGAG 2940

QY 2941 AAAATGCATCAGTCAATATGACCCCGAATTCAAGGAATTTACATTGTAAGAGAAAAC 3000

Db 2941 AAAATGCATCAGTCAATATGACCCCGAATTCAAGGAATTTACATTGTAAGAGAAAAC 3000

QY 3001 AAAGCACTGCAAAAGAAAGTGTGCCGACTATCAAAATAAATGGTGAATCATCTGCAAAATGT 3060

Db 3001 AAAGCACTGCAAAAGAAAGTGTGCCGACTATCAAAATAAATGGTGAATCATCTGCAAAATGT 3060

QY 3061 GGCCAGGCTTGGGGAACAATGATGGTGCAAAAGGCTTAGATTTCCTTGTCTCAAAATA 3120

Db 3061 GGCCAGGCTTGGGGAACAATGATGGTGCAAAAGGCTTAGATTTCCTTGTCTCAAAATA 3120

QY 3121 AGGAATTTTGTAGTGGTTTTCAAAAATAATTCAAACAAAGAAACAATACAAAAAGTGGGTA 3180

Db 3121 AGGAATTTTGTAGTGGTTTTCAAAAATAATTCAAACAAAGAAACAATACAAAAAGTGGGTA 3180

QY 3181 GAATTACCTATCACATTTCCCAATCTTGAAGATTTTAAAAATACTATCAGTTAAACATTTAATATGATTA 3300

Db 3181 GAATTACCTATCACATTTCCCAATCTTGAAGATTTTAAAAATACTATCAGTTAAACATTTAATATGATTA 3300

QY 3301 TGATTAATGTATTCAATTATGCTACAGAACTGACATAAGAAATCAATAAATGATTGTTT 3360

Db 3301 TGATTAATGTATTCAATTATGCTACAGAACTGACATAAGAAATCAATAAATGATTGTTT 3360

QY 3361 CTCTG 3365

Db 3361 CTCTG 3365

RESULT 5
ABA04908

ID ABA04908 standard; cDNA; 3372 BP.

XX ABA04908;

DT 05-MAR-2002 (first entry)

XX Human RNA helicase RH116 coding sequence.

DE Human; RH116; RNA helicase; cytostatic; virucide; anti-HIV;
XX immunosuppressive; immunostimulatory; antirheumatic; antiarthritic;
KW antiarteriosclerotic; osteopathic; antidiabetic; hepatotropic;
KW antiinflammatory; cancer; infection; HIV; hepatitis; genetic disease;
KW autoimmune disease; graft rejection; vaccine; ss.

OS Homo sapiens.

XX Key Location/Qualifiers
CDS 155..3232

FT /*tag= a

FT /product= "Human RH116"

XX WO200185955-A1.

XX 15-NOV-2001.

XX 11-MAY-2001; 2001WO-FR001441.

XX 11-MAY-2000; 2000FR-00006030.

XX (ISTA-) ISTAC.

XX (INSP) INST PASTEUR LILLE.

PI Bahr G, Cocude C, Capron A;

XX WPI; 2002-082898/11.

DR P-PSDB; AAM47798.

XX New polypeptide, useful for treating and diagnosing cancer or
PT inflammation, and drug screening, comprises a human polynucleotide
PT homologous to RNA helicase.

PS	Claim 7; Page 85-89; 114pp; French.
XX	The present sequence is the coding sequence for human RH16. RH16 is a
CC	116kDa protein and has homology to RNA helicases (DEXH box). RH16 and
CC	its coding sequence are useful for treating cancer; acute or chronic
CC	infections (especially by HIV or hepatitis B or C); inherited genetic
CC	diseases; (auto)immune diseases (particularly rheumatism, arthritis,
CC	arteriosclerosis, osteoporosis and diabetes, but many others listed) and
CC	to prevent graft rejection. RH16 and its coding sequence are also useful
CC	for inducing, or increasing, the immune response to a vaccine
XX	
SQ	Sequence 3372 BP; 1157 A; 637 C; 748 G; 830 T; 0 U; 0 Other;
	Query Match 99.2%; Score 3338.8; DB 6; Length 3372;
	Best Local Similarity 99.8%; Pred. No. 0;
	Matches 3343; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY	15 GAGCCCTGTGGACAACTCGTCATTGTGAGGACAGAGCGGTAGACCCTGCTCTCTAAG 74
DB	1 GGGCCCTGTGGACAACTCGTCATTGTGAGGACAGAGCGGTAGACCCTGCTCTCTAAG 60
QY	75 TGGGACGGGACAGCGGCACGCACATTTACCTGTGCCGACAGCAACAGCACCATCTGCT 134
DB	61 TGGGACGGGACAGCGGCACGCACATTTACCTGTGCCGACAGCAACAGCACCATCTGCT 120
QY	135 TGGGAGAACCCCTCTCCCTCTCTGTGAGAAAGAAAGATGTCGAATGGGTATTTCCACAGACGA 194
DB	121 TGGGAGAACCCCTCTCCCTCTCTGTGAGAAAGAAAGATGTCGAATGGGTATTTCCACAGACGA 180
QY	195 GAATTTCCGCTATCTCATCTCGTCTTCAGGSCCAGGGTGAAATGTACATCCAGGTGA 254
DB	181 GAATTTCCGCTATCTCATCTCGTCTTCAGGSCCAGGGTGAAATGTACATCCAGGTGA 240
QY	255 GCCTGTGCTGGACTACCTGACCTTTCTGCCTGCAGAGGTGAAGGACAGATTTCAGAGGAC 314
DB	241 GCCTGTGCTGGACTACCTGACCTTTCTGCCTGCAGAGGTGAAGGACAGATTTCAGAGGAC 300
QY	315 AGTCGCCACCTCCGGGAACATGCAGGCAGTTGAACTGCTGAGCACCTTGGAGAGGG 374
DB	301 AGTCGCCACCTCCGGGAACATGCAGGCAGTTGAACTGCTGAGCACCTTGGAGAGGG 360
QY	375 AGTCGGCACCTTGGTTGGACTCGGGAATTCGTGGAGGCCCTCCGGAGAACCGGCAGCCC 434
DB	361 AGTCGGCACCTTGGTTGGACTCGGGAATTCGTGGAGGCCCTCCGGAGAACCGGCAGCCC 420
QY	435 TCTGGCCGCCGCTACATGAACCTTGAGCTCACGGACTTGCCCTCTCCATCGTTTGAGAA 494
DB	421 TCTGGCCGCCGCTACATGAACCTTGAGCTCACGGACTTGCCCTCTCCATCGTTTGAGAA 480
QY	495 CGTCTATGATGAATATCTCCAACTGCTGAACCTCTCAGCCCCACTCTGTGGACAAGCT 554
DB	481 CGTCTATGATGAATATCTCCAACTGCTGAACCTCTCAGCCCCACTCTGTGGACAAGCT 540
QY	555 TCTAGTTAGAGACGTCTTGGATAAGTGATGGAGGAACTGTTGACAAATTGAAGACAG 614
DB	541 TCTAGTTAGAGACGTCTTGGATAAGTGATGGAGGAACTGTTGACAAATTGAAGACAG 600
QY	615 AAACCGGATTGCTGCTGCAGAAAACAAATGGAATGAATCAGGTGTAAGAGAGCTACTAAA 674
DB	601 AAACCGGATTGCTGCTGCAGAAAACAAATGGAATGAATCAGGTGTAAGAGAGCTACTAAA 660
QY	675 AAGGATTGTGAGAAAAGAAAACCTGGTTCTGCAATTTCTGAATGTTCTTCGTCAAACAGG 734
DB	661 AAGGATTGTGAGAAAAGAAAACCTGGTTCTGCAATTTCTGAATGTTCTTCGTCAAACAGG 720
QY	735 AAACAATGAACCTTGTCCAAGAGTTAACAGGCTCTGATTGCTCAGAAAGCAATGCAGAGAT 794
DB	721 AAACAATGAACCTTGTCCAAGAGTTAACAGGCTCTGATTGCTCAGAAAGCAATGCAGAGAT 780
QY	795 TGAGAAATTTATCACAAAGTTGATGGTCTCAAGTGAAGAGCAACTCTTTCAACACAGT 854
DB	781 TGAGAAATTTATCACAAAGTTGATGGTCTCAAGTGAAGAGCAACTCTTTCAACACAGT 840

QY	855 TCAGCCAAATCTGGAGAGGAGGTCTGGGCGATGGAGAAATAACTCATCAGAAATCATCTTT 914
DB	841 TCAGCCAAATCTGGAGAGGAGGTCTGGGCGATGGAGAAATAACTCATCAGAAATCATCTTT 900
QY	915 TGCAGATTCTTCTGTAGTTTCAGAAATCAGACACAAGTTTGGCAGAAAGAGTGTGAGCTG 974
DB	901 TGCAGATTCTTCTGTAGTTTCAGAAATCAGACACAAGTTTGGCAGAAAGAGTGTGAGCTG 960
QY	975 CTTAGATGAAAGTCTTGGACATAAACAGCAACATGGGCAGTGATTTCAGGCACCATGGGAAG 1034
DB	961 CTTAGATGAAAGTCTTGGACATAAACAGCAACATGGGCAGTGATTTCAGGCACCATGGGAAG 1020
QY	1035 TGATTTCAGATGAAGAGAAATGTGGCAGCAAGAGCATCCCGAGCCAGAACTCCAGCTCAG 1094
DB	1021 TGATTTCAGATGAAGAGAAATGTGGCAGCAAGAGCATCCCGAGCCAGAACTCCAGCTCAG 1080
QY	1095 GCCTTACCCTTGGAGTTGCCCGAGCCAGCCCTTGGAGGGAAAGAAATATCATCATCTGCCT 1154
DB	1081 GCCTTACCCTTGGAGTTGCCCGAGCCAGCCCTTGGAGGGAAAGAAATATCATCATCTGCCT 1140
QY	1155 CCCTACAGGGAGTGGAAAAACCCAGAGTGGCTGTTTACATTGCCAAGGATCACTTAGACAA 1214
DB	1141 CCCTACAGGGAGTGGAAAAACCCAGAGTGGCTGTTTACATTGCCAAGGATCACTTAGACAA 1200
QY	1215 GAAAGAAAAAGCATCTGAGCCTGGAAAAAGTTATAGTTCTTGTCAATAAGGTACTGCTAGT 1274
DB	1201 GAAAGAAAAAGCATCTGAGCCTGGAAAAAGTTATAGTTCTTGTCAATAAGGTACTGCTAGT 1260
QY	1275 TGAACAGCTCTTCGCAAGGAGTTCCAAACCATTTTGAAGAAATGGTATCGTGTATTGG 1334
DB	1261 TGAACAGCTCTTCGCAAGGAGTTCCAAACCATTTTGAAGAAATGGTATCGTGTATTGG 1320
QY	1335 ATTAAGTGGTGATACCCAACTGAAATATCATTTCCAGAAAGTTGTCAAGTCCCTGTGATAT 1394
DB	1321 ATTAAGTGGTGATACCCAACTGAAATATCATTTCCAGAAAGTTGTCAAGTCCCTGTGATAT 1380
QY	1395 TATTATCAGTACAGCTCAAAATCCTTTGAAAACTCCCTCTTAAACTTGGAAAAATGGAGAAGA 1454
DB	1381 TATTATCAGTACAGCTCAAAATCCTTTGAAAACTCCCTCTTAAACTTGGAAAAATGGAGAAGA 1440
QY	1455 TGCTGGTGTTCAAATTTGTGAGACTTTTCCCTCATTTATCATTTGATGAATGTCAACACCAA 1514
DB	1441 TGCTGGTGTTCAAATTTGTGAGACTTTTCCCTCATTTATCATTTGATGAATGTCAACACCAA 1500
QY	1515 CAAAGAACGAGTGATATAATAACATCATGAGGCATTTTGTGATGAGAAAGTTGAAAAACAA 1574
DB	1501 CAAAGAACGAGTGATATAATAACATCATGAGGCATTTTGTGATGAGAAAGTTGAAAAACAA 1560
QY	1575 TAGACTCAAGAAAGAAACAAACCAAGCAGTGATTCCTTCCCTCCTCAGATACTGGGACTAACAGC 1634
DB	1561 TAGACTCAAGAAAGAAACAAACCAAGCAGTGATTCCTTCCCTCCTCAGATACTGGGACTAACAGC 1620
QY	1635 TTCACCTGCTGTTGGAGGGGCCACGAAGCAAGCCAAAGCTGAAGAACACATTTTAAAACT 1694
DB	1621 TTCACCTGCTGTTGGAGGGGCCACGAAGCAAGCCAAAGCTGAAGAACACATTTTAAAACT 1680
QY	1695 ATGTGCCAATCTTGATGCAATTTACTATTATAAACTGTAAAGAAACCTTTGATCAACTGAA 1754
DB	1681 ATGTGCCAATCTTGATGCAATTTACTATTATAAACTGTAAAGAAACCTTTGATCAACTGAA 1740
QY	1755 AAACCAAAATACAGGAGCCATGCAAGAAAGTTTGCATTTGCAGATGCAACACAGAGAAGATCC 1814
DB	1741 AAACCAAAATACAGGAGCCATGCAAGAAAGTTTGCATTTGCAGATGCAACACAGAGAAGATCC 1800
QY	1815 ATTTAAAGAGAAACTTCTAGAAATAATGACAAAGGATTCAAACTTATTGTCAATGAGTCC 1874
DB	1801 ATTTAAAGAGAAACTTCTAGAAATAATGACAAAGGATTCAAACTTATTGTCAATGAGTCC 1860
QY	1875 AATGTGAGATTTTGGAACTCAACCTTATGAACAATGGCCATTCAAAATGGAAGAAAAAGC 1934
DB	1861 AATGTGAGATTTTGGAACTCAACCTTATGAACAATGGCCATTCAAAATGGAAGAAAAAGC 1920
QY	1935 TGCAAAAAAGGAAATCGCAAGAAACGCTGTTTGTGAGAAACATTTGAGGAAAGTACAATGA 1994

||||| 1921 TGCAAAAGGAAGGAAATCGCAAGAAAGTGTGTGCAGAACATTTGAGGAAGTACAATAA 1980

QY 1995 GGCCCTACAAATTAAATGACACAAATTCGAATGATAGATGCGTATATCTCATCTTGAAACTTT 2054

Db 1981 GGCCCTACAAATTAAATGACACAAATTCGAATGATAGATGCGTATATCTCATCTTGAAACTTT 2040

QY 2055 CTATAATGAAGAGAAAGATAAGAAGTTTGCAGTCATAGAAAGATGATAGTGAGGGTGG 2114

Db 2041 CTATAATGAAGAGAAAGATAAGAAGTTTGCAGTCATAGAAAGATGATAGTGAGGGTGG 2100

QY 2115 TGATGATGAGTATTTGATGGTGATGAAGATGAGGATGATTTAAAGAAACCTTTGAAACT 2174

Db 2101 TGATGATGAGTATTTGATGGTGATGAAGATGAGGATGATTTAAAGAAACCTTTGAAACT 2160

QY 2175 GGATGAAACAGATAGATTTCTCATGACTTTATTTTTTGAAACAAATAAATAATGTTGAAAG 2234

Db 2161 GGATGAAACAGATAGATTTCTCATGACTTTATTTTTTGAAACAAATAAATAATGTTGAAAG 2220

QY 2235 GCTGGCTGAAAAACCCAGAAATATGAAAATGAAAAGCTGACCAAAATTAAGAAATACCATAT 2294

Db 2221 GCTGGCTGAAAAACCCAGAAATATGAAAATGAAAAGCTGACCAAAATTAAGAAATACCATAT 2280

QY 2295 GGAGCAATATACTAGGACTGAGGAATCAGCAGCAGGAATAATCTTTACAAAAACACGACA 2354

Db 2281 GGAGCAATATACTAGGACTGAGGAATCAGCAGCAGGAATAATCTTTACAAAAACACGACA 2340

QY 2355 GAGTGCATATGCGCTTTCCAGTGGATTACTGAAAAATGAAAAATTTGCTGAAGTAGGAGT 2414

Db 2341 GAGTGCATATGCGCTTTCCAGTGGATTACTGAAAAATGAAAAATTTGCTGAAGTAGGAGT 2400

QY 2415 CAAAGCCCACCATCTGATTTGGAGCTGGACACAGCAGTGAGTTCAAACCCCATGACACAGAA 2474

Db 2401 CAAAGCCCACCATCTGATTTGGAGCTGGACACAGCAGTGAGTTCAAACCCCATGACACAGAA 2460

QY 2475 TGAACAAAAAGAGTCATTAGTAAATTTGCGACTGGAAAAATCAATCTGCTTATCGCTAC 2534

Db 2461 TGAACAAAAAGAGTCATTAGTAAATTTGCGACTGGAAAAATCAATCTGCTTATCGCTAC 2520

QY 2535 CACAGTGGCAGAGAGGTCTGGATATTTAAAGAATGTAACATTTGTTATCCGTTATGGTCT 2594

Db 2521 CACAGTGGCAGAGAGGTCTGGATATTTAAAGAATGTAACATTTGTTATCCGTTATGGTCT 2580

QY 2595 CGTCACCAATGAATAGCCATGGTCCAGGCCCGTGGTCGAGCCAGAGCTGATGAGAGCAC 2654

Db 2581 CGTCACCAATGAATAGCCATGGTCCAGGCCCGTGGTCGAGCCAGAGCTGATGAGAGCAC 2640

QY 2655 CTACGTCCTGGTTGCTCACAGTGGTTCAGAGTTATCGAACATGAGACAGTTAATGATTT 2714

Db 2641 CTACGTCCTGGTTGCTCACAGTGGTTCAGAGTTATCGAACATGAGACAGTTAATGATTT 2700

QY 2715 CCGAGAGAAGATGATGTATAAAGCTATACATTTGTGTTCAAAATATGAAACCAGAGGAGTA 2774

Db 2701 CCGAGAGAAGATGATGTATAAAGCTATACATTTGTGTTCAAAATATGAAACCAGAGGAGTA 2760

QY 2775 TGCTCATAAGATTTTGGAAATTACAGATGCAAAAGTATAATGGAAGAAAAATGAAAACCAA 2834

Db 2761 TGCTCATAAGATTTTGGAAATTACAGATGCAAAAGTATAATGGAAGAAAAATGAAAACCAA 2820

QY 2835 GAGAAATATTGCCAAGCATTTACAAGAAATAACCCATCACTAATAACTTTCCTTTGCAAAA 2894

Db 2821 GAGAAATATTGCCAAGCATTTACAAGAAATAACCCATCACTAATAACTTTCCTTTGCAAAA 2880

QY 2895 CTGCAGTGTGCTAGCCTGTTCTGGGGAAGATATCCATGTAAATTGAGAAAAATGCATCACGT 2954

Db 2881 CTGCAGTGTGCTAGCCTGTTCTGGGGAAGATATCCATGTAAATTGAGAAAAATGCATCACGT 2940

QY 2955 CAATATGACCCCCAGAAATCAAGGAACTTTACATTTGTAAGAGAAAAACAAAGCACTGC AAAA 3014

Db 2941 CAATATGACCCCCAGAAATCAAGGAACTTTACATTTGTAAGAGAAAAACAAAGCACTGC AAAA 3000

QY 3015 GAAAGTGCCGACTATCAAAATAAATGGTGAATCATCTGCAAAATGTGCCAGGCTTGGGG 3074

Db 3001 GAAGTGTGCCGACTATCAAATAAAATGGTGAAATCATCTGCAAAATGTGGCCAGGCTTGGG 3060

QY 3075 AACAAATGATGGTGCAAAAAGGCTTAGATTTGCCTTGTCTCTCAAATAAGGAATTTTGTAGT 3134

Db 3061 AACAAATGATGGTGCAAAAAGGCTTAGATTTGCCTTGTCTCTCAAATAAGGAATTTTGTAGT 3120

QY 3135 GGTTTTCAAAAATAAATTCACAAAGAAAAACAATACAAAAAGTGGGTAGAAATTACCTATCAC 3194

Db 3121 GGTTTTCAAAAATAAATTCACAAAGAAAAACAATACAAAAAGTGGGTAGAAATTACCTATCAC 3180

QY 3195 ATTTCCCAATCTTGACTATTTCAGAAATGCTGTTTATTTTAGTGATGAGGATTAGCACTTGAT 3254

Db 3181 ATTTCCCAATCTTGACTATTTCAGAAATGCTGTTTATTTTAGTGATGAGGATTAGCACTTGAT 3240

QY 3255 TGAAGATTTCTTTAAATACTATCAGTTAAACATTTTAATATGATTATGATTATGATTTC 3314

Db 3241 TGAAGATTTCTTTAAATACTATCAGTTAAACATTTTAATATGATTATGATTATGATTTC 3300

QY 3315 ATTATGCTACAGAACTGACATAAGAAATCAATAAAATGATTGTTTTACTCT 3364

Db 3301 ATTATGCTACAGAACTGACATAAGAAATCAATAAAATGATTGTTTTACTCT 3350

RESULT 6

ADC30823

ID ADC30823 standard; cDNA; 3446 BP.

XX

AC ADC30823;

XX

DT 18-DEC-2003 (first entry)

XX

DE Human novel cDNA sequence, SEQ ID NO:905.

XX

KW Human; diagnostic; drug screening; forensics; gene mapping;

KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;

KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;

KW ulcers; osteoporosis; autoimmune disease; cancer;

KW molecular weight marker; food supplement; antiparkinsonian; nootropic;

KW neuroprotective; antianaemic; anticoagulant; thrombolytic; vulnerary;

KW antiulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;

KW gene therapy; gene; ss.

XX

OS Homo sapiens.

XX

PN WO2003029271-A2.

XX

PD 10-APR-2003.

XX

PF 24-SEP-2002; 2002WO-US030474.

XX

PR 24-SEP-2001; 2001US-0324631P.

XX

PA (HYSE-) HYSEQ INC.

XX

PI Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;

PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;

PI Haley-Vicente D, Drmanac RT;

XX

DR WPI; 2003-371981/35.

DR P-PSDB; ADC31794.

XX

PT New polynucleotide and polypeptide useful for diagnosing, preventing or

PT treating conditions such as neurodegenerative diseases, anemias, platelet

PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or

PT cancer.

XX

PS Claim 1; SEQ ID NO 905; 1185pp; English.

XX

CC The invention relates to 971 novel human cDNA sequences (ADC29919-

CC ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The

CC invention also relates to nucleic acid sequences over 99% identical with

CC the novel human cDNAs. The invention additionally encompasses expression

CC vectors and host cells comprising a nucleic acid of the invention; the

CC recombinant production of a polypeptide of the invention; an antibody
CC against a polypeptide of the invention; a method of detecting
CC polynucleotides or polypeptides of the invention; and methods of
CC identifying a compound which binds to a polypeptide of the invention. The
CC invention further discloses methods of preventing, treating or
CC ameliorating a medical condition; kits comprising polynucleotide probes
CC and/or monoclonal antibodies for carrying out the methods of the
CC invention; methods for the identification of compounds that modulate the
CC expression or activity of the polynucleotide and/or polypeptide; and 767
CC contig sequences corresponding to the cDNA sequences of the invention
CC (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628
CC -ADC33394). The nucleic acids and polypeptides of the invention are
CC useful in diagnostics, drug screening, forensics, gene mapping, in the
CC identification of mutations responsible for genetic disorders or other
CC traits, for assessing biodiversity, and in producing many other types of
CC data and products dependent on DNA and amino acid sequences. They are
CC also used for treating diseases such as Parkinson's disease, Alzheimer's
CC disease and other neurodegenerative diseases, anaemia, platelet
CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
CC cancer. The nucleic acids may also be used as hybridisation probes or
CC primers, and in the recombinant production of a protein. The polypeptides
CC are also useful in generating antibodies, as molecular weight markers,
CC and as food supplements. The present sequence represents a specifically
CC claimed human cDNA sequence of the invention. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX
SQ Sequence 3446 BP; 1184 A; 649 C; 772 G; 841 T; 0 U; 0 Other;

Query Match 98.5%; Score 3316.2; DB 10; Length 3446;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 3321; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY	36	CATTGTCAGGCACAGAGCGGTAGACCCCTGCTTCTTAAGTGGCAGCGGACAGCGCACG	95
Db	104	CACCTTCAGGCACAGAGCGGTAGACCCCTGCTTCTTAAGTGGCAGCGGACAGCGCACG	163
QY	96	CACATTTACCTGTCCCGCAGACAAACAGCACCATCTGCTTGGGAGAACCTCTCCCTTCT	155
Db	164	CACATTTACCTGTCCCGCAGACAAACAGCACCATCTGCTTGGGAGAACCTCTCCCTTCT	223
QY	156	CTGAGAAAGAAAGATGTCGAATGGGTATTCACAGACGAGAAATTCGGCTATCTCATCTC	215
Db	224	CTGAGAAAGAAAGATGTCGAATGGGTATTCACAGACGAGAAATTCGGCTATCTCATCTC	283
QY	216	GTGCTTCAGGGCCAGGGTGAATGTAATCCAGTGGAGCCCTGTGTGGACTACCTGAC	275
Db	284	GTGCTTCAGGGCCAGGGTGAATGTAATCCAGTGGAGCCCTGTGTGGACTACCTGAC	343
QY	276	CTTCTGCTGCAGAGGTGAAGAGCAGATTTCAGAGGACAGTCGCCACCTCCGGGAACAT	335
Db	344	CTTCTGCTGCAGAGGTGAAGAGCAGATTTCAGAGGACAGTCGCCACCTCCGGGAACAT	403
QY	336	GCAGGCAGTTGAACCTGCTGCTGAGCACCTTGGAGAGGGAGTCTGGCACTTGGTTGGAC	395
Db	404	GCAGGCAGTTGAACCTGCTGCTGAGCACCTTGGAGAGGGAGTCTGGCACTTGGTTGGAC	463
QY	396	TCGGGAATTCCTGGAGGCCCTCCGAGAACCCGGCAGCCCTCTGGCCCGCGCTACATGAA	455
Db	464	TCGGGAATTCCTGGAGGCCCTCCGAGAACCCGGCAGCCCTCTGGCCCGCGCTACATGAA	523
QY	456	CCCTGAGCTCAGGACTTGGCCCTCTCCATCGTTTGAGAACGCTCATGATGAATATCTCCA	515
Db	524	CCCTGAGCTCAGGACTTGGCCCTCTCCATCGTTTGAGAACGCTCATGATGAATATCTCCA	583
QY	516	ACTGCTGAACCTCTTCAGCCCACTCTGGTGGACAAGCTTCTAGTTAGAGACGCTCTTGA	575
Db	584	ACTGCTGAACCTCTTCAGCCCACTCTGGTGGACAAGCTTCTAGTTAGAGACGCTCTTGA	643
QY	576	TAAGTGCATGGAGGAACTGTTGACAAATTGAAGACAGAAACCGGATTGCTGCTGCAGA	635
Db	644	TAAGTGCATGGAGGAACTGTTGACAAATTGAAGACAGAAACCGGATTGCTGCTGCAGA	703

QY	636	AAACAATGGAAATGAATCAGGTGTAAGAGAGCTACTAAAAAGGATTGTGCAGAAAGAAA	695
Db	704	AAACAATGGAAATGAATCAGGTGTAAGAGAGCTACTAAAAAGGATTGTGCAGAAAGAAA	763
QY	696	CTGGTTCTCTGCATTTCTGAATGTTCTTCGTCAAAACAGGAAACAAATGTTGTCCAAGA	755
Db	764	CTGGTTCTCTGCATTTCTGAATGTTCTTCGTCAAAACAGGAAACAAATGTTGTCCAAGA	823
QY	756	GTTAACAGGCTCTGATTGCTCAGAAAGCAATGCAGAGATTGAGAAATTTATCACAAGTTGA	815
Db	824	GTTAACAGGCTCTGATTGCTCAGAAAGCAATGCAGAGATTGAGAAATTTATCACAAGTTGA	883
QY	816	TGGTCTCTCAAGTGAAGAGCAACTTCTTTCAACCCACAGTTTCAGCCAAATCTGGAGAAAG	875
Db	884	TGGTCTCTCAAGTGAAGAGCAACTTCTTTCAACCCACAGTTTCAGCCAAATCTGGAGAAAG	943
QY	876	GGTCTGGGCGATGGAGAAATAACTCATCAGAAATCATCTTTTCAGATTCTTCTGTAGTTTC	935
Db	944	GGTCTGGGCGATGGAGAAATAACTCATCAGAAATCATCTTTTCAGATTCTTCTGTAGTTTC	1003
QY	936	AGAATCAGACACAAGTTTGGCAGAGGAAGTGTCTCAGTCTCTTAGATGAAAGTCTTGGACA	995
Db	1004	AGAATCAGACACAAGTTTGGCAGAGGAAGTGTCTCAGTCTCTTAGATGAAAGTCTTGGACA	1063
QY	996	TAAACAGCAACATGGGCGAGTGAATTCAGGCACCATGGGAAGTATTTCAGATGAAGAGAAATGT	1055
Db	1064	TAAACAGCAACATGGGCGAGTGAATTCAGGCACCATGGGAAGTATTTCAGATGAAGAGAAATGT	1123
QY	1056	GGCAGCAAGAGCATCCCGGAGCCAGAACTCCAGCTCAGGCCTTACCAATGGAAGTTGC	1115
Db	1124	GGCAGCAAGAGCATCCCGGAGCCAGAACTCCAGCTCAGGCCTTACCAATGGAAGTTGC	1183
QY	1116	CCAGCCAGCCTTGGAAAGGGAAGATATCATCATCTGCTCCCTACAGGGAGTGGAAAAAC	1175
Db	1184	CCAGCCAGCCTTGGAAAGGGAAGATATCATCATCTGCTCCCTACAGGGAGTGGAAAAAC	1243
QY	1176	CAGAGTGGCTGTTTACATTGCCAAGGATCACTTAGACAAGAAAAAAGCATCTGAGCC	1235
Db	1244	CAGAGTGGCTGTTTACATTGCCAAGGATCACTTAGACAAGAAAAAAGCATCTGAGCC	1303
QY	1236	TGAAAAAGTTATAGTCTTGTCAATAAGTACTGCTAGTTGAACAGCTCTTCCGCAAGGA	1295
Db	1304	TGAAAAAGTTATAGTCTTGTCAATAAGTACTGCTAGTTGAACAGCTCTTCCGCAAGGA	1363
QY	1296	GTTCCAAACCATTTTGAAGAAATGGTATCGTGTATTGGATTAAAGTGGTGATACCCAACT	1355
Db	1364	GTTCCAAACCATTTTGAAGAAATGGTATCGTGTATTGGATTAAAGTGGTGATACCCAACT	1423
QY	1356	GAAATATCATTTCCAGAAAGTTGTCAAGTCTCTGTGATATTATTCAGTACAGCTCAAT	1415
Db	1424	GAAATATCATTTCCAGAAAGTTGTCAAGTCTCTGTGATATTATTCAGTACAGCTCAAT	1483
QY	1416	CCTTGAAAACTCCCTCTTAAACTTGGAAAAATGGAGAAAGATGCTGGTGTCAATTGTGAGA	1475
Db	1484	CCTTGAAAACTCCCTCTTAAACTTGGAAAAATGGAGAAAGATGCTGGTGTCAATTGTGAGA	1543
QY	1476	CTTTTCCCTCATTTATCATTTGATGAATGTCTATCACCAACAAAGAGCAGTGTATAATAA	1535
Db	1544	CTTTTCCCTCATTTATCATTTGATGAATGTCTATCACCAACAAAGAGCAGTGTATAATAA	1603
QY	1536	CATCATGAGGCATTTATTTGATGCAGAAAGTTGAAAAACAATAGACTCAAGAAAAAACA	1595
Db	1604	CATCATGAGGCATTTATTTGATGCAGAAAGTTGAAAAACAATAGACTCAAGAAAAAACA	1663
QY	1596	ACCAGTATTCCTTCCCTCAGATACCTGGGACTAAGCAGCTTACCTGGTGTGGAGGGGC	1655
Db	1664	ACCAGTATTCCTTCCCTCAGATACCTGGGACTAAGCAGCTTACCTGGTGTGGAGGGGC	1723
QY	1656	CACGAAGCAAGCCAAAGCTGAAGAACACATTTTAAACTATGTGCCAATCTTGATGCATT	1715
Db	1724	CACGAAGCAAGCCAAAGCTGAAGAACACATTTTAAACTATGTGCCAATCTTGATGCATT	1783

QY 1716 TACTATTAAAACTGTTAAAGAAAAACCTTGATCAACTGAAAAACCAAATACAGGAGCCATG 1775
Db |
1784 TACTATTAAAACTGTTAAAGAAAAACCTTGATCAACTGAAAAACCAAATACAGGAGCCATG 1843
QY 1776 CAAGAAAGTTTGCCATTGCAGATGCAACCCAGAGAAGATCCATTAAAGAGAAACTTCTAGA 1835
Db |
1844 CAAGAAAGTTTGCCATTGCAGATGCAACCCAGAGAAGATCCATTAAAGAGAAACTTCTAGA 1903
QY 1836 AATAATGACAAGGATTCAAACTTATTGTCAAAATGAGTCCAATGTGAGATTTTGGAACTCA 1895
Db |
1904 AATAATGACAAGGATTCAAACTTATTGTCAAAATGAGTCCAATGTGAGATTTTGGAACTCA 1963
QY 1896 ACCCTATGAACAATGGGCCATTCAAAATGGBAAAAAAGCTGCAAAAAAAGGAAATCGCAA 1955
Db |
1964 ACCCTATGAACAATGGGCCATTCAAAATGGBAAAAAAGCTGCAAAAAAAGGAAATCGCAA 2023
QY 1956 AGAACGTGTTTGTGCAGAACATTTGAGGAAGTACAAATGAGGCCCTACAAATTAATGACAC 2015
Db |
2024 AGAAAGTGTGTCAGAACATTTGAGGAAGTACAAATAGGCCCTACAAATTAATGACAC 2083
QY 2016 AATTGGAATGATAGATGCGTATACTCATCTTGAAAACTTCTATAATGAAGAGAAAAATAA 2075
Db |
2084 AATTGGAATGATAGATGCGTATACTCATCTTGAAAACTTCTATAATGAAGAGAAAAATAA 2143
QY 2076 GAAAGTTGCAGTCATAGAAGATGATAGTGAGGGTGGTGATGATGAGTATTGTGATGG 2135
Db |
2144 GAAAGTTGCAGTCATAGAAGATGATAGTGAGGGTGGTGATGATGAGTATTGTGATGG 2203
QY 2136 TGATGAAGATGAGGATGATTTAAAGAAACCTTTGAAACTGGATGAAACAGATAGATTTCT 2195
Db |
2204 TGATGAAGATGAGGATGATTTAAAGAAACCTTTGAAACTGGATGAAACAGATAGATTTCT 2263
QY 2196 CATGACTTTATTTTGTAAAAACAATAAATGTTGAAAAAGCTGGTGAAAAACCCAGATA 2255
Db |
2264 CATGACTTTATTTTGTAAAAACAATAAATGTTGAAAAAGCTGGTGAAAAACCCAGATA 2323
QY 2256 TGAAAAATGAAAAGCTGACCAAAATTAAGAAATACCATAATGGAGCAATATACTAGGACTGA 2315
Db |
2324 TGAAAAATGAAAAGCTGACCAAAATTAAGAAATACCATAATGGAGCAATATACTAGGACTGA 2383
QY 2316 GGAATCAGCACGAGGAATAATCTTTACAAAAACACGACAGAGTGATATGCGCTTCCCA 2375
Db |
2384 GGAATCAGCACGAGGAATAATCTTTACAAAAACACGACAGAGTGATATGCGCTTCCCA 2443
QY 2376 GTGGATTACTGAAAATGAAAAATTTGCTGAAGTAGGAGTCAAGCCACCATCTGATTGG 2435
Db |
2444 GTGGATTACTGAAAATGAAAAATTTGCTGAAGTAGGAGTCAAGCCACCATCTGATTGG 2503
QY 2436 AGCTGGACACAGCAGTGAGTTCAAACCCCATGACACAGAATGAACAAAAAGAAGTCATTAG 2495
Db |
2504 AGCTGGACACAGCAGTGAGTTCAAACCCCATGACACAGAATGAACAAAAAGAAGTCATTAG 2563
QY 2496 TAAATTTGCGACTGAAAAAATCAATCTGCTTATCGCTACCACAGTGGCAGAGAAGGTCT 2555
Db |
2564 TAAATTTGCGACTGAAAAAATAAATCTGCTTATCGCTACCACAGTGGCAGAGAAGGTCT 2623
QY 2556 GGATATTAAAGAAATGTAAACATTGTTATCCGTTATGGTCTCGTCACCAATGAAATAGCCAT 2615
Db |
2624 GGATATTAAAGAAATGTAAACATTGTTATCCGTTATGGTCTCGTCACCAATGAAATAGCCAT 2683
QY 2616 GGTCCAGGCCGTGGTCGAGCCAGAGCTGATGAGAGCACCTACGTCCTGGTTGCTCACAG 2675
Db |
2684 GGTCCAGGCCGTGGTCGAGCCAGAGCTGATGAGAGCACCTACGTCCTGGTTGCTCACAG 2743
QY 2676 TGGTTCAGGAGTTATCGAACATGAGACAGTTAATGATTTCCGAGAGAAGATGATGTATAA 2735
Db |
2744 TGGTTCAGGAGTTATCGAACCTGAGACAGTTAATGATTTCCGAGAGAAGATGATGTATAA 2803
QY 2736 AGCTATACATTGTGTTCAAAAATATGAAACACAGAGGAGTATGCTCATAGATTTTGGAAAT 2795
Db |
2804 AGCTATACATTGTGTTCAAAAATATGAAACACAGAGGAGTATGCTCATAGATTTTGGAAAT 2863
QY 2796 ACAGATGCAAAGTATAATGGAAAAAGAAAATGAAAAACCAAGAGAAATATTGCCAAGCATTA 2855

Db |
2864 ACAGATGCAAAGTATAATGGAAAAAGAAAATGAAAAACCAAGAGAAATATTGCCAAGCATTA 2923
QY 2856 CAAGAAATAACCCATCACTAATAACTTTCTTTGCAAAAACTGCAGTGTGCTAGCCTGTTTC 2915
Db |
2924 CAAGAAATAACCCATCACTAATAACTTTCTTTGCAAAAACTGCAGTGTGCTAGCCTGTTTC 2983
QY 2916 TGGGGAAGATATCCATGTAAATTGAGAAAAATGCATCACGTCAATATGACCCCGAGAAATCAA 2975
Db |
2984 TGGGGAAGATATCCATGTAAATTGAGAAAAATGCATCACGTCAATATGACCCCGAGAAATCAA 3043
QY 2976 GGAACTTTACATTTGAAGAGAAAAACAAGCAGCTGCAAAAGAGTGTGCCGACTATCAAAAT 3035
Db |
3044 GGAACTTTACATTTGAAGAGAAAAACAAGCAGCTGCAAAAGAGTGTGCCGACTATCAAAAT 3103
QY 3036 AAATGGTGAATCATCTGCAAAATGTGGCCAGGCTTGGGGAACAATGATGGTGCAAAAGG 3095
Db |
3104 AAATGGTGAATCATCTGCAAAATGTGGCCAGGCTTGGGGAACAATGATGGTGCAAAAGG 3163
QY 3096 CTTAGATTTGCCTTGTCTCAAAATAAGGAATTTTGTAGTGGTTTTCAAAAAATAATTCAAC 3155
Db |
3164 CTTAGATTTGCCTTGTCTCAAAATAAGGAATTTTGTAGTGGTTTTCAAAAAATAATTCAAC 3223
QY 3156 AAAGAAACAATACAAAAAGTGGGTAGAATTACCTATCACATTTCCCAATCTTGACTATTTC 3215
Db |
3224 AAAGAAACAATACAAAAAGTGGGTAGAATTACCTATCACATTTCCCAATCTTGACTATTTC 3283
QY 3216 AGAATGCTGTTTATTAGTGATGAGGATTAGCACCTTGATTGAAGATTCTTTTAAATAACT 3275
Db |
3284 AGAATGCTGTTTATTAGTGATGAGGATTAGCACCTTGATTGAAGATTCTTTTAAATAACT 3343
QY 3276 ATCAGTTAAACATTTAATATGATTATGATTAAATGATTTCATTATGCTACAGAACTGCAT 3335
Db |
3344 ATCAGTTAAACATTTAATATGATTATGATTAAATGATTTCATTATGCTACAGAACTGCAT 3403
QY 3336 AAGAATCAATAAAATGATTGTTTTACTCT 3364
Db |
3404 AAGAATCAATAAAATGATTGTTTTACTCT 3432

RESULT 7

AAD11170

ID AAD11170 standard; DNA; 3131 BP.

XX AAD11170;

XX 29-NOV-2001 (first entry)

XX Human melanoma differentiation associated-5 protein-related DNA.

XX Human melanoma differentiation associated gene; Mda-5; interferon; IFN;
RNA helicase motif; caspase recruitment domain; CARD; therapy; melanoma;
neuroblastoma; astrocytoma; glioblastoma; multiforme; cancer; cervical;
breast; colon; prostate; osteosarcoma; chondrosarcoma; systemic toxicity;
central nervous system; cytosstatic; apoptosis; ds.

OS Homo sapiens.

XX WO200164707-A1.

XX 07-SEP-2001.

XX 28-FEB-2001; 2001WO-US0006960.

PR 29-FEB-2000; 2000US-00515363.

XX (UYCO) UNIV COLUMBIA NEW YORK.

XX Fisher PB, Kang D, Gopalkrishnan RV;

XX WPI; 2001-565494/63.

PT Nucleic acid sequences encoding a Melanoma Differentiation Associated

PT Gene useful for cancer cell growth suppression, apoptosis and anti-viral
PT activity.

XX Disclosure; Page 134-148; 152pp; English.

XX The present invention relates to an isolated nucleic acid encoding a
CC melanoma differentiation associated gene-5 (Mda-5) polypeptide. Mda-5
CC contains a caspase recruitment domain (CARD) and a RNA helicase motif.
CC Mda-5 is a novel interferon (IFN) inducible gene with structural
CC similarities to RNA helicases and CARD motif containing proteins. Mda-5
CC is induced during terminal differentiation in human melanoma cells
CC treated with the combination of recombinant fibroblast IFN and the
CC antileukaemic compound mezerein (MEZ). Mda-5 is useful for identifying
CC compounds that may induce its expression. Mda-5 is useful for treating
CC cancer e.g., melanoma, neuroblastoma, astrocytoma, glioblastoma
CC multiforme, cervical cancer, breast cancer, colon cancer, prostate
CC cancer, osteosarcoma, chondrosarcoma and especially where the cancer is a
CC cancer of the central nervous system and apoptosis. The Mda-5 promoter
CC exhibits melanocyte tissue specificity and minimises systemic toxicity.
CC The present sequence is human Mda-5 protein-related DNA. Note: The
CC present sequence is designated SEQ ID NO:2 in the sequence listing, but
CC does not correspond to the sequence designated SEQ ID NO:2 in the main
CC body of the specification (AAE10155). The present sequence is not further
CC referred to in the specification, and has been represented in a protein
CC format in the sequence listing

XX Sequence 3131 BP; 1075 A; 583 C; 699 G; 774 T; 0 U; 0 Other;

Query Match 92.2%; Score 3104; DB 4; Length 3131;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 3129; Conservative 0; Mismatches 0; Indels 5; Gaps 2;

QY 169 ATGTCGAATGGGTATCCACAGACGAGAAATTCGGCTATCTCATCTCGTCTTCAAGGCC 228
DB 1 ATGTCGAATGGGTATCCACAGACGAGAAATTCGGCTATCTCATCTCGTCTTCAAGGCC 60
QY 229 AGGGTGAATAATGATACATCCAGGTGGAGCTGTGTGGACTACCTGACCTTTCTGCTGCA 288
DB 61 AGGGTGAATAATGATACATCCAGGTGGAGCTGTGTGGACTACCTGACCTTTCTGCTGCA 120
QY 289 GAGGTGAAGGAGCAGATTCCAGGACAGTCGCCACCTCCGGGAACATGCAGGCAGTTGAA 348
DB 121 GAGGTGAAGGAGCAGATTCCAGGACAGTCGCCACCTCCGGGAACATGCAGGCAGTTGAA 180
QY 349 CTGCTGCTGAGCACCTTGGAGAAAGGAGTCTGGCACCTTGGTGGACTCGGGAATTCGTG 408
DB 181 CTGCTGCTGAGCACCTTGGAGAAAGGAGTCTGGCACCTTGGTGGACTCGGGAATTCGTG 240
QY 409 GAGGCCCTCCGGAGAACCGGCAGCCCTCTGGCCGCCCGCTACATGAACCTTGAGCTCACG 468
DB 241 GAGGCCCTCCGGAGAACCGGCAGCCCTCTGGCCGCCCGCTACATGAACCTTGAGCTCACG 300
QY 469 GACTTGCCCTCTCCATCGTTTGAGAACCGTCAATGATGAATATCTCCAACTGCTGAACCTC 528
DB 301 GACTTGCCCTCTCCATCGTTTGAGAACCGTCAATGATGAATATCTCCAACTGCTGAACCTC 360
QY 529 CTTGAGCCCACTCTGTGGACAAGCTTCTAGTTAGAGACGCTTGGATAAGTGCATGGAG 588
DB 361 CTTGAGCCCACTCTGTGGACAAGCTTCTAGTTAGAGACGCTTGGATAAGTGCATGGAG 420
QY 589 GAGGAACCTGTTGACAAATTGAAGACAGAAACCGGATTTGCTGTCGAGAAAACCAATGGAAT 648
DB 421 GAGGAACCTGTTGACAAATTGAAGACAGAAACCGGATTTGCTGTCGAGAAAACCAATGGAAT 480
QY 649 GAATCAGGTGTAAGAGAGCTACTAAAAGGATTTGTCAGAAAAGAAACTGTTCTCTGCA 708
DB 481 GAATCAGGTGTAAGAGAGCTACTAAAAGGATTTGTCAGAAAAGAAACTGTTCTCTGCA 540
QY 709 TTTCTGAATGTTCTTCTCAACAGGAAACCAATGAACTTGTCCAAAGTTAACAGGCTCT 768
DB 541 TTTCTGAATGTTCTTCTCAACAGGAAACCAATGAACTTGTCCAAAGTTAACAGGCTCT 600
QY 769 GATTGCTCAGAAAGCAATGCAGAGATTGAGAAATTTATCACAAGTTGATGGTCTCAAGTG 828

DB 601 GATTGCTCAGAAAGCAATGCAGAGATTGAGAAATTTATCACAAGTTGATGGTCTCAAGTG 660
QY 829 GAAGAGCAACTTCTTTCAACCACACAGTTCAGCCAAATCTGGAGAGGAGGTCTGGGGCATG 888
DB 661 GAAGAGCAACTTCTTTCAACCACACAGTTCAGCCAAATCTGGAGAGGAGGTCTGGGGCATG 720
QY 889 GAGAATAACTCATCAGAAATCATCTTTTGAGATTCTTCTGTAGTTTTCAGAAATCAGACACA 948
DB 721 GAGAATAACTCATCAGAAATCATCTTTTGAGATTCTTCTGTAGTTTTCAGAAATCAGACACA 780
QY 949 AGTTTGGCAGAAAGGAGTGTCTAGCTGCTTAGATGAAAGTCTTTGGACATAAACAGCAACATG 1008
DB 781 AGTTTGGCAGAAAGGAGTGTCTAGCTGCTTAGATGAAAGTCTTTGGACATAAACAGCAACATG 840
QY 1009 GGCAGTGATTTCAGGCACCATGGGAAGTGATTTCAGATGAAGAGAAATGTGGCAGCAAGAGCA 1068
DB 841 GGCAGTGATTTCAGGCACCATGGGAAGTGATTTCAGATGAAGAGAAATGTGGCAGCAAGAGCA 900
QY 1069 TCCCCGGAGCCAGAACTCCAGCTCAGGCCTTACCAATGGAAGTTGCCCGCCAGCCCTTG 1128
DB 901 TCCCCGGAGCCAGAACTCCAGCTCAGGCCTTACCAATGGAAGTTGCCCGCCAGCCCTTG 960
QY 1129 GAAGGGAAGAAATATCATCATCTGCTTCCCTACAGGGAGTGGAAAACACAGAGTGGCTGTT 1188
DB 961 GAAGGGAAGAAATATCATCATCTGCTTCCCTACAGGGAGTGGAAAACACAGAGTGGCTGTT 1020
QY 1189 TACATTGCCAAGGATCACTTAGACAAAGAAAAGAGCATCTGAGCTGGAAAAGTTATA 1248
DB 1021 TACATTGCCAAGGATCACTTAGACAAAGAAAAGAGCATCTGAGCTGGAAAAGTTATA 1080
QY 1249 GTTCTTGTCAATAAGGTACTGTAGTTGAACAGACTCTTCCGCAAGGAGTTCCAACCATTT 1308
DB 1081 GTTCTTGTCAATAAGGTACTGTAGTTGAACAGACTCTTCCGCAAGGAGTTCCAACCATTT 1140
QY 1309 TTGAAGAAATGGTATCGTGTATTGGATTAAAGTGGTGATACCCAACTGAAAATATCATTT 1368
DB 1141 TTGAAGAAATGGTATCGTGTATTGGATTAAAGTGGTGATACCCAACTGAAAATATCATTT 1200
QY 1369 CCAGAAAGTTGTCAAGTCTGTGTATATTATTATCAGTACAGCTCAAAATCCTTGAAAACCTC 1428
DB 1201 CCAGAAAGTTGTCAAGTCTGTGA--TATTATCAGTACAGCTCAAAATCCTTGAAAACCTC 1257
QY 1429 CTCTTAAACTTGGAAAATGGAGAAAGATGCTGGTGTTCATTTGTCAGACTTTTCCCTCATTT 1488
DB 1258 CTCTTAAACTTGGAAAATGGAGAAAGATGCTGGTGTTCATTTGTCAGACTTTTCCCTCATTT 1317
QY 1489 ATCATTTGATGAATGTATCATCACCAACAAAGAAAGCAGTGTATAATAATCATCATGAGGCAT 1548
DB 1318 ATCATTTGATGAATGTATCATCACCAACAAAGAAAGCAGTGTATAATAATCATCATGAGGCAT 1377
QY 1549 TATTTGATGCAGAAAGTTGAAAAACAATAGACTCAAGAAAAGAAAACAAACAGTGATTTCCC 1608
DB 1378 TATTTGATGCAGAAAGTTGAAAAACAATAGACTCAAGAAAAGAAAACAAACAGTGATTTCCC 1437
QY 1609 CTTCTCTCAGATACCTGGGACTTAACAGCTTCACCTGGTGTGGAGGGGCCAGGAAGCAAGCC 1668
DB 1438 CTTCTCTCAGATACCTGGGACTTAACAGCTTCACCTGGTGTGGAGGGGCCAGGAAGCAAGCC 1497
QY 1669 AAAGCTGAAGAACACATTTTAAAAACTATGTGCCAATCTTGTATGCTATTTACTATTAAAACT 1728
DB 1498 AAAGCTGAAGAACACATTTTAAAAACTATGTGCCAATCTTGTATGCTATTTACTATTAAAACT 1557
QY 1729 GTTAAAGAAAACCTTTGATCAACTGAAAAACCAATACAGGAGCCATGCAAGAGTTTGCC 1788
DB 1558 GTTAAAGAAAACCTTTGATCAACTGAAAAACCAATACAGGAGCCATGCAAGAGTTTGCC 1617
QY 1789 ATTGCAGATGCAACCAGAGAGATCCATTTTAAAGAGAAAACCTTCTAGAAATTAATGACAAGG 1848
DB 1618 ATTGCAGATGCAACCAGAGAGATCCATTTTAAAGAGAAAACCTTCTAGAAATTAATGACAAGG 1677
QY 1849 ATTCAAACCTTATTTGTCAAATGAGTCCAAATGTAGATTTTGGAACTCAACCCCTATGAACAA 1908

Db 1678 ATTCAAACTTATTGTCAAATGAGTCCAATGTGAGATTTTGGAACTCAACCCCTATGAACAA 1737

Qy 1909 TGGGCCATTCAAATGGAAGAAAAAGCTGCAAAAAAGGAAATCGCAAGAAACGTGTTGT 1968

Db 1738 TGGGCCATTCAAATGGAAGAAAAAGCTGCAAAAAAGGAAATCGCAAGAAACGTGTTGT 1797

Qy 1969 GCAGAAACATTGAGGAAGTACAATGAGGCCCTACAAATTAATGACACAATTCGAATGATA 2028

Db 1798 GCAGAAACATTGAGGAAGTACAATGAGGCCCTACAAATTAATGACACAATTCGAATGATA 1857

Qy 2029 GAT - GCGTATACTCATCTTGAAACTTTCTATAATGAAGAGAAAGATAAGATTTCGAG 2086

Db 1858 GATCCGCGTATACTCATCTTGAAACTTTCTATAATGAAGAGAAAGATAAGATTTCGAG 1917

Qy 2087 TCATAGAAAGATGATAGTATGAGGGTGGTGATGAGTATTGTGATGGTGATGAAGATG 2146

Db 1918 TCATAGAAAGATGATAGTATGAGGGTGGTGATGAGTATTGTGATGGTGATGAAGATG 1977

Qy 2147 AGGATGATTTAAAGAAACCTTTGAAACTGGATGAAACAGATAGATTCTCATGACTTTAT 2206

Db 1978 AGGATGATTTAAAGAAACCTTTGAACTGGATGAAACAGATAGATTCTCATGACTTTAT 2037

Qy 2207 TTTTGTGAAAAACAATAAATGTTGAAAAGCTGGCTGAAAAACCCAGAAATATGAAAAATGAAA 2266

Db 2038 TTTTGTGAAAAACAATAAATGTTGAAAAGCTGGCTGAAAAACCCAGAAATATGAAAAATGAAA 2097

Qy 2267 AGCTGACCAAAATTAAGAAATACCATAAATGGAGCAATATACTAGGACTGAGGAATCAGCAC 2326

Db 2098 AGCTGACCAAAATTAAGAAATACCATAAATGGAGCAATATACTAGGACTGAGGAATCAGCAC 2157

Qy 2327 GAGGAATAATCTTTTACAAAAACACGACAGAGTGCATATGCGCTTTCCAGTGGATTACTG 2386

Db 2158 GAGGAATAATCTTTTACAAAAACACGACAGAGTGCATATGCGCTTTCCAGTGGATTACTG 2217

Qy 2387 AAAATGAAAAATTTGCTGAAGTAGGAGTCAAAGCCCACCATCTGATTGGAGCTGGACACA 2446

Db 2218 AAAATGAAAAATTTGCTGAAGTAGGAGTCAAAGCCCACCATCTGATTGGAGCTGGACACA 2277

Qy 2447 GCAGTGAGTTCAAACCCATGACACAGAAATGAACAAAAAGAGTCAATAGTAAATTTTCGCA 2506

Db 2278 GCAGTGAGTTCAAACCCATGACACAGAAATGAACAAAAAGAGTCAATAGTAAATTTTCGCA 2337

Qy 2507 CTGGAATAATCAATCTGCTTATCGTCTACACAGTGGCAGAGAGTCTGGATATTAAG 2566

Db 2338 CTGGAATAATCAATCTGCTTATCGTCTACACAGTGGCAGAGAGTCTGGATATTAAG 2397

Qy 2567 AATGTAACATTGTTATCCGTTATGGTCTCGTCACCAATGAAATAGCCATGGTCCAGGCC 2626

Db 2398 AATGTAACATTGTTATCCGTTATGGTCTCGTCACCAATGAAATAGCCATGGTCCAGGCC 2457

Qy 2627 GTGGTCGAGCCAGAGCTGATGAGAGCACCTACGTCCTGGTTGCTACAGTGGTTCAGGAG 2686

Db 2458 GTGGTCGAGCCAGAGCTGATGAGAGCACCTACGTCCTGGTTGCTACAGTGGTTCAGGAG 2517

Qy 2687 TTATCGAAACATGAGACAGTTAATGATTTCCGAGAGAAGATGATGATATAAGCTATACATT 2746

Db 2518 TTATCGAAACATGAGACAGTTAATGATTTCCGAGAGAAGATGATGATATAAGCTATACATT 2577

Qy 2747 GTGTTCAAAATATGAACACAGAGGAGTATGCTCATAAGATTTTGGAAATACAGATGCAAA 2806

Db 2578 GTGTTCAAAATATGAACACAGAGGAGTATGCTCATAAGATTTTGGAAATACAGATGCAAA 2637

Qy 2807 GTATAATGGAAGAAATGAAAAACCAAGAGAAATATTGCCAAGCATTTACAAGATAACC 2866

Db 2638 GTATAATGGAAGAAATGAAAAACCAAGAGAAATATTGCCAAGCATTTACAAGATAACC 2697

Qy 2867 CATCACTAATAACTTTTCCGAACTGAGTGTGCTAGCCTGTTCTGGGGAAGATA 2926

Db 2698 CATCACTAATAACTTTTCCGAACTGAGTGTGCTAGCCTGTTCTGGGGAAGATA 2757

Qy 2927 TCCATGTAATTGAGAAATGCATCACGTCAATATGACCCAGAAATTCAGGAACTTTTACA 2986

Db 2758 TCCATGTAATTGAGAAATGCATCACGTCAATATGACCCAGAAATTCAGGAACTTTTACA 2817

Qy 2987 TTGTAAGAGAAAAACAAAGCACACTGCAAAAAGAAAGTGTCCGACTATCAAAATAAATGGTGAAA 3046

Db 2818 TTGTAAGAGAAAAACAAAGCACACTGCAAAAAGAAAGTGTCCGACTATCAAAATAAATGGTGAAA 2877

Qy 3047 TCATCTGCAAAATGTGGCCAGGCTTGGGGAACAATGATGGTGCAAAAGGCTTAGATTGTC 3106

Db 2878 TCATCTGCAAAATGTGGCCAGGCTTGGGGAACAATGATGGTGCAAAAGGCTTAGATTGTC 2937

Qy 3107 CTTGTCTCAAAATAAGGAATTTTGTAGTGGTTTTTCAAAAATAAATCAACAAAGAAACAAT 3166

Db 2938 CTTGTCTCAAAATAAGGAATTTTGTAGTGGTTTTTCAAAAATAAATCAACAAAGAAACAAT 2997

Qy 3167 ACAAAAAGTGGTAGAATTACCTATCACATTTCCCAATCTTGACTATTTCAGAAATGCTGTT 3226

Db 2998 ACAAAAAGTGGTAGAATTACCTATCACATTTCCCAATCTTGACTATTTCAGAAATGCTGTT 3057

Qy 3227 TATTTAGTGATGAGGATTAGCACCTTGATTGAAGATTCTTTTAAAAATACTATCAGTTAAAC 3286

Db 3058 TATTTAGTGATGAGGATTAGCACCTTGATTGAAGATTCTTTTAAAAATACTATCAGTTAAAC 3117

Qy 3287 ATTTAATATGATTA 3300

Db 3118 ATTTAATATGATTA 3131

RESULT 8

ADJ75813

ID ADJ75813 standard; DNA; 3771 BP.

XX

AC ADJ75813;

XX

DT 20-MAY-2004 (first entry)

XX

DE Marker gene SEQ ID NO:1065.

XX

KW bronchial asthma; chronic obstructive pulmonary disease;

KW respiratory epithelial cell; interleukin-13; respiratory; antiasthmatic;

KW gene therapy; marker gene; gene; ds.

XX

OS Mus musculus.

XX

PN EP1394274-A2.

XX

PD 03-MAR-2004.

XX

PF 04-AUG-2003; 2003EP-00254857.

XX

PR 06-AUG-2002; 2002JP-00229312.

PR

20-MAR-2003; 2003JP-00077212.

XX

PA (GENO-) GENOX RES INC.

XX

PI Ohtani N, Sugita Y, Yamaya M, Kubo H, Nagai H, Izuvara K;

XX

DR WPI; 2004-193155/19.

XX

PT Testing for bronchial asthma or chronic obstructive pulmonary disease by comparing the expression level of a marker gene in a biological sample from a subject with the expression level of the gene in a sample from a healthy subject.

PS Claim 14; SEQ ID NO 1065; 241pp; English.

XX

CC The present invention describes a method of testing for bronchial asthma or chronic obstructive pulmonary disease. The method comprises determining the expression level of a marker gene in a biological sample from a subject, comparing the expression level determined with the expression level of the marker gene in a biological sample from a healthy subject, and judging whether the subject has bronchial asthma or chronic obstructive pulmonary disease. The marker gene comprises: (a) a group of genes (S1) whose expression levels increase when respiratory epithelial cells are stimulated with interleukin-13; or (b) a group of genes (S2)

CC whose expression levels decrease when respiratory epithelial cells are
CC stimulated with interleukin-13. Also described: (1) a reagent (I) for
CC testing for bronchial asthma or chronic obstructive pulmonary disease;
CC (2) a kit for screening for a candidate compound for a therapeutic agent
CC to treat bronchial asthma or chronic obstructive pulmonary disease; (3)
CC an animal model for bronchial asthma or chronic obstructive pulmonary
CC disease; (4) an inducer that induces bronchial asthma in a mouse; (5) a
CC method for producing an animal model for bronchial asthma or chronic
CC obstructive pulmonary disease; (6) a therapeutic agent for bronchial
CC asthma or chronic obstructive pulmonary disease, comprising the compound,
CC a marker gene or an antisense nucleic acid corresponding to a portion of
CC the marker gene, a ribozyme, a polynucleotide that suppresses the
CC expression of the gene through an RNAi effect or an antibody recognising
CC a protein encoded by a marker gene; and (7) a DNA chip for testing for
CC bronchial asthma or a chronic obstructive pulmonary disease, on which a
CC probe has been immobilised to assay a marker gene. (I) has respiratory
CC and antiasthmatic activities, and can be used in gene therapy. The method
CC is useful for testing for or screening for a therapeutic agent for
CC bronchial asthma or chronic obstructive pulmonary disease. The present
CC sequence is used in the exemplification of the present invention.

XX
SQ Sequence 3771 BP; 1183 A; 821 C; 866 G; 901 T; 0 U; 0 Other;
Query Match 65.5%; Score 2202.4; DB 12; Length 3771;
Best Local Similarity 80.7%; Pred. No. 0;
Matches 2609; Conservative 0; Mismatches 616; Indels 7; Gaps 3;

QY	121	CAGCACCATCTGCTGGGAGAACCTCTCC-TTCTCTGAGAAAGAAAGATGTCGAATGG	179
DB	186	CGCGCCCGCCGCCGGGAGACTCCTCTCCCATTTCTGAGACCTCAGCATGTCGATTGT	245
QY	180	GTATTCCACAGACGAGAAATTTCCGCTATCTCATCTCGTGTTCAGGGCCAGGGTGAATAAT	239
DB	246	CTGTTCTGCAGAGGACAGCTTCAGGAATCTCATCTTATTTTCAGGCCCAGGCTGAATAAT	305
QY	240	GTACATCCAGGTGGAGCCTGTGCTGGACTACCTGACCTTTCTGCTGCAGAGGTGAAGGA	299
DB	306	GTACATTCAGGTGGAGCCAGTGTGGACCACTCTTCTGTCTGCAGAAACCAAGA	365
QY	300	GCAGATTTCAGAGGACAGTCGCCACCTCCGGGAACATGCAGGCAGTTGAACTGCTGCTGAG	359
DB	366	GCAGATTCTTAAAAAGATCAACACCTGTGGTAACACACAGCGCGGCAAGAACTGCTGCTGAG	425
QY	360	CACCTTGGAGAGGGAGTCTGGCACCTTGGTTGGACTCGGGAATTCGTGGAGGCCCTCCG	419
DB	426	CACCTTGGAGCAGGACAATGGCCTCTGGGATGGACGCAGATGTTCTGGAGGCCCTAGA	485
QY	420	GAGAACCGGCAGCCCTCTGGCCGCCCTGCTGACATGACACCTTGAGTCAACCTTGCCCTC	479
DB	486	GCACAGTGGCAATCCCTTAGCCGCGCTATGTCAAAACCCACACTCACTGATCTGCCCTC	545
QY	480	TCCATCGTTGAGAACGCTCATGATGAATATCTCCAACTGCTGAACCTCTTCAGCCCCAC	539
DB	546	TCCTTCTCTGAGACTGCCCATGACGAGTGTCTCCACTTGTGACCTCTCCAGCCCCAC	605
QY	540	TCTGGTGGACAAGCTTCTAGTTAGACAGCTCTTGGATAAGTGCAATGGAGGGAAGTGT	599
DB	606	TTTGGTGGACAACCTTCTGATTAAACGATGCTTGGACACTTGTTCGAGAAAGGACTATT	665
QY	600	GACAATTGAGACAGAAACCGGATTGCTGTCAGAAACAATGGAATGAATCAGGTGT	659
DB	666	GACAGTCGAGACAGAAATCGGATTCTGCTGCAGGAAACAGCGGGAATGAGTCAGGTGT	725
QY	660	AAGAGAGCTACTAAAAAGGATTGTGCAGAAAGAAACTGGTTCTCTGCAATTTCTGAATGT	719
DB	726	AAGAGAGCTGCTGAGAAAGGATTGTGCAGAAAGGAAACTGGTTTCTTACTCTCTGGATGT	785
QY	720	TCTTCGTCAAAACAGGAAACAATGAATGTCCAAGAGTTAAACAGGCTCTGATGCTCAGA	779
DB	786	TCTGCGCCAACTGGAAATGATGCACATTTCCAAGAACTAACAGGTGGAGGCTGCCCAGA	845
QY	780	AAGCAATGCAGAGATTGAGAATTTATCAAAAGTTGATGGTCTCAAGTGGAGAGCAACT	839

DB	846	AGACAAACACAGACTTGGCTAACTCGTCTCACAGAGATGGCCCTGCAGCTAATGAGTGTCT	905
QY	840	TCCTTCAACCACAGTTTCAGCCAAATCTGGAGAAAGAGGCTCTGGGCGATGGAGATAACTC	899
DB	906	TCTGCCTGCTGTCGATGAGTCAAGTCTGGAGACAGAGGCCCTGGAACGTAGACGACATATT	965
QY	900	ATCAGAAATCATCTTTTTCAGATCTTCTGTGATTTTCAGAAATCAGACACAAAGTTTGGCAGA	959
DB	966	ACCAGAGGCTTCTTGTACAGATTCTGTGTGACCACAGAAATCAGACACAAAGTTTGGCAGA	1025
QY	960	AGGAAGTGTGAGCTGCTTAGATGAAAGTCTTGGACATAACAGCAACATGGGCAGTATTC	1019
DB	1026	AGGAAGTGTGAGCTGCTTTCGATGAAAGTCTTGGACATAACAGCAACATGGGCAGGATTC	1085
QY	1020	AGGCACCATGGGAAAGTGAATTCAGATGAA--GAGAAATGTGGCAGCAAGAGCATCCCCGA	1076
DB	1086	AGGCACCATGGGAAAGTGAATTCAGATGAAAGTGTCTATCCAGACAAAAGAGTATCCCCGA	1145
QY	1077	GCCAGAACTCCAGCTCAGGCCTTACCAAAATGGAAGTTGCCAGCCAGCCTTGGAAAGGAA	1136
DB	1146	GCCAGAACTGCAGCTCAGGCCTTACCAAAATGGAAGTGGCCCAACAGCTCTAGATGGGAA	1205
QY	1137	GAATATCATCTGCTCCCTACAGGGAGTGGAAAAACAGAGTGGCTGTTTACATTC	1196
DB	1206	GAATATTTATCTGCTCCCTCCCGGAGTGGGAAAACAGAGTGGCTGTTTACATCAC	1265
QY	1197	CAAGGATCATTTAGACAAGAAAAAGCATCTGAGCCTGGAAAAAGTTATAGTCTTGT	1256
DB	1266	CAAAGATCATTTAGACAAGAAAGCAGGCATCTGAATCCGGGAAGGTTATCGTCTTGT	1325
QY	1257	CAATAAGGTACTGCTAGTTGAACAGCTCTTCCGCAAGGAGTTCCAACCATTTTGAAGAA	1316
DB	1326	CAATAAGGTAATGTTAGCAGAACAACTTTCCGAAAAAGAGTTCAACCCATATTTGAAGAA	1385
QY	1317	ATGGTATCGTGTATTGGATTAAGTGGATACCCCACTGAAAAATATCATTTCCAGAAAT	1376
DB	1386	ATGGTATCGAATTTATGGATTAAGTGGGATACCCAGCTGAAAAATATCATTTCCAGAAAT	1445
QY	1377	TGTCAAGTCTGTGATATTATTATCAGTACAGCTCAAATCTTGAATACTCCCTCTTAAA	1436
DB	1446	TGTCAAAATCTTACGATGTTATTATCAGCACTGCTCAAATCTTGAATACTCCCTCTTAAA	1505
QY	1437	CTTGGAAAAATGGAGAGATGCTGGTTCATTTGTGAGCTTTTCCCTCATTTATCATTTGA	1496
DB	1506	TCTGGAGAGTGGAGACGATGACGGTGTGAGCTGTGAGCTTCTCTCATTTATCATTTGA	1565
QY	1497	TGAATGTCTATCACACCAACAAAGAGCAGTGTATATAACATCATGAGGCATTTATTTGAT	1556
DB	1566	TGAGTGCCTATCACACCAACAAAGGAGGCGAGTCTATAACACATCATGAGACGATATTTGAA	1625
QY	1557	GCAGAAAGTTGAAAAACAATFAGACTCAAGAAAGAAAAACAAACAGTATTTCCCTTCTCA	1616
DB	1626	GCAGAAAGTGGAGAAACAATGACCTCAAGAAACAAAAACAAACAGCCATTTCCCTGCGCA	1685
QY	1617	GATACTGGGACTAACAGCTTTCACCTGGTGTGGAGGGGCCACGAAAGCAAGCCAAAGCTGA	1676
DB	1686	GATACTAGGACTGACAGCTTTCACCTGGTGTGGAGGAGCCAAAAAGCAGTCTGAGGCTGA	1745
QY	1677	AGAACACATTTTAAAACTATGTGCCAATCTTGTGATGCAATTTACTATTAATACTGTTAAAGA	1736
DB	1746	AAACATATTTTAAATATATGTGCCAATCTTGTGATGCTTTTACCATTAAAAACAGTGAAGA	1805
QY	1737	AAACCTTGTATCAACTGAAAAACCAAAATACAGGAGCCATGCAAGAAAGTTTGCATTTGCAGA	1796
DB	1806	GAATCTTGGTCAACTCAAAACCAAAATAAAGGAACCATGCAAGAAATTTGTGATTTGCTGA	1865
QY	1797	TGCAACCCAGAGAGATCCATTTTAAAGAGAACTTCTAGAAAAATATGACAAAGGATTTCAAAC	1856
DB	1866	TGACACCCAGAGAAAAATCCATTTTAAAGAGAACTTCTAGAAAAATTTATGGCAAGCATTTAGAC	1925
QY	1857	TTATTGTCAATGAGTCCCAATGTCAGATTTTGGAACTCAACCCCTATGAACAAATGGGCCAT	1916
DB	1926	TTACTGCCAAAAAAGTCCAATGTCAGATTTTGGAAACCCCAACATTTATGAGCAGTGGGCCAT	1985

QY	1917	TCAAATGGA	AAAAAGCTG	CAAAAAA	AGGAAATCG	CAAGAACG	TGTTTGT	GCAGAAC	A	1976	
DB	1986	TCAAATGG	AGAAAAAGCTG	CTAAAGAC	GCGAAATCG	CAAGATCG	CGTCTGT	GCAGAAC	A	2045	
QY	1977	TTTGAGGA	AGTACAA	TGAGGCC	CTACAAATTA	TGACACA	AAATTCGA	ATGATAG	ATGCGTA	2036	
DB	2046	TTTGAGGA	AGTACAA	CGAAGCC	CTACAAATCA	ACGACAC	GGATCCGA	ATGATG	ATGATG	2105	
QY	2037	TACTCAT	CTTTGAA	ACTTTCT	ATAATGA	AGAAAG	ATAAGAA	GTTTGC	AGTCAT	AAGA	2096
DB	2106	TAGCCAC	CTTGAGG	CATTTCA	CATGATG	AGAAAG	AAAGAA	GTTTGC	AGTCCT	CAA	2162
QY	2097	TGATAGT	GATGAGG	TGGTGAT	GATGAT	GATTTGT	GATGAT	GATGAG	GATGAT	TAT	2156
DB	2163	TGACAGC	GACAAAG	AGTGATG	ACGAGG	CCAGCA	GTGCAAT	GACCA	CTTAAGG	CGATGT	2222
QY	2157	AAAGAA	ACCTTTG	AACTGG	ATGAA	ACAGAT	AGATTTCT	CATGAC	TTTATTTT	TGAAA	2216
DB	2223	AAAGAA	ATCTTTG	AACTGG	ACGAA	ACGGAT	GAATTTCT	CATGA	ATTTGTT	TGATA	2282
QY	2217	CAATAAA	ATGTTGA	AAAGGCT	GGCTG	AAACCC	CAGAA	TATGAA	AAATGAA	AGCTGAC	2276
DB	2283	CAAGAA	ATGTTGA	AAAAA	CTAGCTG	AAAA	CCCA	ANATCG	AGATGAA	AAACTCAT	2342
QY	2277	ATTAAGA	AAATCCA	TAAATG	GAGCA	ATATAC	TAGACT	GAGGA	ATCAG	CACGAGGA	2336
DB	2343	ATTAAGA	AAACAC	GATACT	TGAA	CAATTC	CAAGG	CTTG	AGGATCCT	CCCGAGGA	2402
QY	2337	CTTTA	CAAAAA	CACG	ACAGAT	GCATAT	GGCTTT	CCCAG	TGGATTACT	GAAAA	2396
DB	2403	TTTTA	CAAAAA	CACG	ACAGAC	CACTAC	GCACTTT	CCAG	TGGATCAT	GAAAA	2462
QY	2397	ATTTG	CTAAG	TAGAGT	CAAA	GCCCC	CACTCT	GATTTG	GAGCTGG	ACACAGCAGT	2456
DB	2463	GTTT	GCGGA	AGTTG	GAGTCA	AAAGCG	CATCAC	CTGATTTG	CGCGG	GCACAGCAGT	2522
QY	2457	CAAA	CCCATG	ACACAGA	ATGAA	CAAAA	AGAAGT	CATTAG	TAAATTT	CGCACTGGA	2516
DB	2523	CAAG	CCCATG	ACTCAG	ACTGAA	CAAAA	AGAAGT	CATTAG	TAAATTT	CGCACTGGA	2582
QY	2517	CAAT	CTGCTT	ATCGCT	ACCACAG	TGGCAG	AGAAGG	TCTGG	ATATTA	AGAAATGTA	2576
DB	2583	AAAT	CTGCTT	ATCGCT	ACGACGG	TGGCAG	AGAAGG	CTGG	ATATCA	AAAGATGCA	2642
QY	2577	TGTT	ATCCGTT	ATGTT	TCGTCA	CCAATG	AAATAG	CCATGGT	CCAGG	CCCGTGGT	2636
DB	2643	TGTT	ATTCGTT	ATGGCCT	TGTACG	NA	CGAGAT	AGCCAT	GGTCC	AGGCCCGG	2702
QY	2637	CAGAG	CTGATG	AGAGCAC	CTACG	TCTGGT	TGCTCAC	AGTGGT	TTCAGG	AGTTATCG	2696
DB	2703	CAGAG	CTGATG	AAAGCAG	CTATG	TCTGGT	CTACC	AGCAGT	GGCTCAGG	AGTTACCG	2762
QY	2697	TGAG	ACAGTT	TAATGAT	TTCCG	AGAGA	AGATGAT	GTATAA	AGCTATAC	ATTGTTCA	2756
DB	2763	GGAG	ATTGTT	TAATGAT	TTCCG	AGAGA	AGATGAT	GTATAA	AGCTATTA	ACCGTTCA	2822
QY	2757	TATGA	AA	CCAGAGG	AGTATG	CTCAT	AAATTTT	TGAA	TTACAG	ATGCAA	2816
DB	2823	CATGA	AA	CCAGAGG	AGTATG	CA	ATAGATTTT	TGAA	TTTG	AGGATGCA	2882
QY	2817	AAAG	AAAA	ATGAAA	ACCA	AGAA	ATATTG	CCA	AGCATTA	CAAGATTA	2876
DB	2883	AAAG	AAAA	ATGAAA	AGTCA	AAAA	AGCA	ATTG	CAAG	CAATCA	2942
QY	2877	AACT	TTTCT	TTTG	CAAAAA	CTGCAG	TGTG	CTAG	CCCTG	TTCTGGG	2936
DB	2943	AACAC	TTCT	CTGCA	AAAA	TTGTAG	CATGCT	GTCTG	CTCGG	AGAAAA	3002
QY	2937	TGAG	AAAA	ATGC	ATCA	CGTCA	ATATG	ACCC	CAGAA	TTCAAG	2996
DB	3003	TGAG	AGATG	CATCAT	GTCAT	GTCAT	ATGAC	ACC	CAGAA	TTCAAGG	3062

QY	2997	AAACAAAGCACTGCAAAAAGAGTGTGCCGACTATCAAAATAAATGGTGAATCATCTGCAA	3058
Db	3063	AAACAAAGCACTGCAAAAAGAAATTTGCTGATTATCAGACCAATGGAGAGATTATCTGCAA	3122
QY	3057	ATGTGGCCAGGCTTGGGGAACAATGATGGTGCACAAAGGCTTAGATTTCCTTGTCTCAA	3116
Db	3123	GTGTGGCCAGGCTTGGGGAACAATGATGGTGCACAAAGGTTTAGATTTCCTTGTCTTAA	3182
QY	3117	AATAAGGAATTTGTAGTGGTTTTCAAAAATAATTCAACAAAGAAACAATACAAAAAGTG	3176
Db	3183	AATAAGGAATTTGTAGTCAATTTTCAAAAATAAATCACCAGAAACAAGTACAAAGAAGTG	3242
QY	3177	GGTAGAATTACCTATCACATTTCCCAATCTTGACTATTTCAGAAATGCTGTTTATTAGTGA	3236
Db	3243	GGTGAATTGCCTATCAGATTTCCCTGATCTTGACTACTCAGAAATACTGCTTGTATAGTGA	3302
QY	3237	TGAGGATTAGCACTTGATTGAAGATTCTTTTAAATACTATCAGTTAAACATTTAATATG	3296
Db	3303	TGAAGATTAGCACTTGATTATCATGATTATTTTAAATACTGCCAACTCAACATTTGAATATG	3362
QY	3297	ATTATGATTAAATGATTATTCATTATGTACAGAACTGACATAAGAATCAATAAA	3348
Db	3363	ATTTTAATTGTGTTTTCGCTACACTACTGAGCTAACTCTAGAATAATGTAGA	3414
RESULT 9			
AAS40960			
ID	AAS40960 standard; cDNA; 1967 BP.		
XX			
AC	AAS40960;		
XX			
XX	17-DEC-2001 (first entry)		
DE	cDNA encoding novel human enzyme polypeptide #176.		
XX			
KW	Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;		
KW	ligase; hyperproliferative disorder; immunodeficiency disorder;		
KW	autoimmune disorder; neurological disorder; metabolic disorder;		
KW	inflammatory disorder; cardiovascular disorder; reproductive disorder;		
KW	blood-related disorder; infectious disorder; gene therapy; cytostatic;		
KW	anti arthritic; nephrotropic; anticoagulant; ss.		
XX			
OS	Homo sapiens.		
XX			
PN	WO200155301-A2.		
XX			
PD	02-AUG-2001.		
XX			
PF	17-JAN-2001; 2001WO-US001239.		
XX			
PR	31-JAN-2000; 2000US-0179065P.		
PR	04-FEB-2000; 2000US-0180628P.		
PR	24-FEB-2000; 2000US-0184664P.		
PR	02-MAR-2000; 2000US-0186350P.		
PR	16-MAR-2000; 2000US-0189874P.		
PR	17-MAR-2000; 2000US-0190076P.		
PR	18-APR-2000; 2000US-0198123P.		
PR	19-MAY-2000; 2000US-0205515P.		
PR	07-JUN-2000; 2000US-0209467P.		
PR	28-JUN-2000; 2000US-0214886P.		
PR	30-JUN-2000; 2000US-0215135P.		
PR	07-JUL-2000; 2000US-0216647P.		
PR	07-JUL-2000; 2000US-0216880P.		
PR	11-JUL-2000; 2000US-0217487P.		
PR	11-JUL-2000; 2000US-0217496P.		
PR	14-JUL-2000; 2000US-0218290P.		
PR	26-JUL-2000; 2000US-0220963P.		
PR	26-JUL-2000; 2000US-0220964P.		
PR	14-AUG-2000; 2000US-0224518P.		
PR	14-AUG-2000; 2000US-0224519P.		
PR	14-AUG-2000; 2000US-0225213P.		
PR	14-AUG-2000; 2000US-0225214P.		
PR	14-AUG-2000; 2000US-0225266P.		

PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.

PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-02559678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-465566/50.
DR P-PSDB; AAU23090.
XX
PT Novel polypeptides and polynucleotides useful for diagnosing, preventing,
PT treating neural, immune system, muscular, reproductive, pulmonary,
PT cardiovascular, renal, proliferative disorders and cancerous diseases.
XX
PS Claim 4; SEQ ID NO 186; 1180pp; English.
XX
CC The present invention relates to the isolation of novel human enzyme
CC polypeptides (AAU22915-AAU23814), and the cDNA and genomic sequences
CC encoding them. The enzyme polypeptides of the invention may comprise the
CC functional classes of oxidoreductases, transferases, hydrolases, lyases,
CC isomerases or ligases. The sequences of the invention are useful in the
CC diagnosis, treatment, prevention and/or prognosis of a wide range of
CC disorders including hyperproliferative disorders (e.g. cancer),
CC immunodeficiency disorders (e.g. AIDS) autoimmune disorders (e.g.
CC arthritis), neurological disorders (e.g. Alzheimer's disease), metabolic
CC disorders (e.g. phenylketonuria), inflammatory disorders (e.g. asthma),
CC cardiovascular disorders (e.g. atherosclerosis), blood-related disorders
CC (e.g. haemophilia), reproductive disorders (e.g. infertility) and
CC infectious disorders (e.g. Influenza). The polynucleotides of the
CC invention can also be used in gene therapy. AAS40785-AAS41684 represent
CC cDNA sequences encoding for the novel human enzyme polypeptides of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX

SQ	Sequence 1967 BP; 723 A; 344 C; 416 G; 476 T; 0 U; 8 Other;														
	Query Match	55.8%;	Score 1878;	DB 4;	Length 1967;										
	Best Local Similarity	98.5%;	Pred. No. 0;												
	Matches 1941;	Conservative 5;	Mismatches 18;	Indels 7;	Gaps 5;										
QY	1047	AGAGAAATGTGCAGCAAGAGCATCCCCGGAGCCAGAACTCCAGCTCAGGCCCTTACC	AAAT	1106											
Db	2	AGAGAAATGTGCAGCAAGAGCATCCCCGGAGCCAGAACTCCAGCTCAGGCCCTTACC	AAAT	61											
QY	1107	GGAAGTTGCCAGCCAGCCTTGGAAAGGGAAGAATATCATCTCGCTCCCTACAGGGAG	1166												
Db	62	GGAAGTTGCCAGCCAGCCTTGGAAAGGGAAGAATATCATCTCGCTCCCTACAGGGAG	121												
QY	1167	TGGA AAAACCCAGAGTGGCTGTTTACATTGCCAAGGATCACTTAGACAAGAAAAAGC	1226												
Db	122	TGGA AAAACCCAGAGTGGCTGTTTACATTGCCAAGGATCACTTAGACAAGAAAAAGC	181												
QY	1227	ATCTGAGCCTGGAAAAGTTATAGTTCCTTGTCAATAAGGTACTGCTAGTTGAACAGCTCTT	1286												
Db	182	ATCTGAGCCTGGAAAAGTTATAGTTCCTTGTCAATAAGGTACTGCTAGTTGAACAGCTCTT	241												
QY	1287	CCGCAAGGAGTTC CAACCATTTTGAAGAAATGGTATCGTGTATTGGATTAAAGTGTGA	1346												
Db	242	CCGCAAGGAGTTC CAACCATTTTGAAGAAATGGTATCGTGTATTGGATTAAAGTGTGA	301												
QY	1347	TACCCAACTGAAAATATCATTTCCAGAAAGTGTCAAAGTCCCTGTGATATTATTATCAGTAC	1406												
Db	302	TACCCAACTGAAAATATCATTTCCAGAAAGTGTCAAAGTCCCTGTGATATTATTATCAGTAC	361												
QY	1407	AGCTCAAATCCTTGAAAACTCCCTCTTAAACTTGGAAAATGGAGAAGATGCTGGTGTTC	1466												
Db	362	AGCTCAAATCCTTGAAAACTCCCTCTTAAACTTGGAAAATGGAGAAGATGCTGGTGTTC	421												
QY	1467	ATTGTCAGACTTTTCCCTCATTTATCATTTGATGAATGTCAACACCAACAAAGAACGAGT	1526												
Db	422	ATTGTCAGACTTTTCCCTCATTTATCATTTGATGAATGTCAACACCAACAAAGAACGAGT	481												
QY	1527	GTATAATAACATCATGAGGCATTATTTTGATGCAGAAAGTTGAAAACAATAGACTCAAGAA	1586												
Db	482	GTATAATAACATCATGAGGCATTATTTTGATGCAGAAAGTTGAAAACAATAGACTCAAGAA	541												
QY	1587	AGAAAACAAACCACTGATTCCTTCCCTCAGATACTGGGACTTAACAGCTTCACCTGGTGT	1646												
Db	542	AGAAAACAAACCACTGATTCCTTCCCTCAGATACTGGGACTTAACAGCTTCACCTGGTGT	601												
QY	1647	TGGAGGGCCACGAAGCAAGCCAAAGCTGAAGAACACATTTTAAACTATGTGCCAATCT	1706												
Db	602	TGGAGGGCCACGAAGCAAGCCAAAGCTGAAGAACACATTTTAAACTATGTGCCAATCT	661												
QY	1707	TGATGCATTTACTATTAAAACTGTTTAAAGAAAACCTTGATCAACTGAAAAACCAATACA	1766												
Db	662	TGATGCATTTACTATTAAAACTGTTTAAAGAAAACCTTGATCAACTGAAAAACCAATACA	721												
QY	1767	GGAGCCATGCAAGAAAGTTTGCCATTGCGAGATGCAACCCAGAGAAGATCCATTAAAGAGAA	1826												
Db	722	GGAGCCATGCAAGAAAGTTTGCCATTGCGAGATGCAACCCAGAGAAGATCCATTAAAGAGAA	781												
QY	1827	ACTTCTAGAAAATAATGACAAGGATTCAAACTTATTGTCAAATGAGTCCAATGTCAGATT	1886												
Db	782	ACTTCTAGAAAATAATGACAAGGATTCAAACTTATTGTCAAATGAGTCCAATGTCAGATT	841												
QY	1887	TGGAACCTCAACCCATGAACAATGGGCCATTCAAAATGGAAAAAAAAGCTGCAAAAAAGG	1946												
Db	842	TGGAACCTCAACCCATGAACAATGGGCCATTCAAAATGGAAAAAAAAGCTGCAAAAAAGG	901												
QY	1947	AAATCGCAAGAACGTTTGTGCAGAACATTTGAGGAAGTACAAATGAGGCCCTACAAAT	2006												
Db	902	AAATCGCAAGAACGTTTGTGCAGAACATTTGAGGAAGTACAAATGAGGCCCTACAAAT	961												
QY	2007	TAATGACACAATTCGAATGATAGATGCGTATCTCTTGAAACTTTCTATAATGAAGA	2066												
Db	962	TAATGACACAATTCGAATGATAGATGCGTATCTCTTGAAACTTTCTATAATGAAGA	1021												

Qy	2067	GAAAGATAAGAAAGTTTGCAGTCTATAGAAGATGATAGTATGATGAGGGTGGTATGATGAGTA	2126
Db	1022	GAAAGATAAGAAAGTTTGCAGTCTATAGAAGATGATAGTATGATGAGGGTGGTATGATGAGTA	1081
Qy	2127	TTGTGATGGTGATGAAGATGAGGATGATTTTAAAGAAACCTTTGAAAACCTGGATGAAACAGA	2186
Db	1082	TTGTGATGGTGATGAAGATGAGGATGATTTTAAAGAAACCTTTGAAAACCTGGATGAAACAGA	1141
Qy	2187	TAGATTTCTCATGACTTTATTTTTTGA AAAACAATAAAATGTTGAAAAGGCTGGCTGAAAA	2246
Db	1142	TAGATTTCTCATGACTTTATTTTTTGA AAAACAATAAAATGTTGAAAAGGCTGGCTGAAAA	1201
Qy	2247	CCCAGAAATGAAAAATGAAAAGCTGACCAAAATTAAGAAAATACCATAATGGAGCAATATAC	2306
Db	1202	CCCAGAAATGAAAAATGAAAAGCTGACCAAAATTAAGAAAATACCATAATGGAGCAATATAC	1261
Qy	2307	TAGGACTGAGGAATCAGCACGAGGAATAATCTTTACAAAACACGACAGAGTGCATATGC	2366
Db	1262	TAGGACTGAGGAATCAGCACGAGGAATAATCTTTACAAAACACGACAGAGTGCATATGC	1321
Qy	2367	GCTTTCCAGTGGATTACTGAAAATGAAAATTTTGCTGAAGTAGGAGTCAAAGCCCCACCA	2426
Db	1322	GCTTTCCAGTGGATTACTGAAAATGAAAATTTTGCTGAAGTAGGAGTCAAAGCCCCACCA	1381
Qy	2427	TCTGATTGGAGCTGGACACAGCAGTGAGTTCAAAACCCATGACACAGAAATGAACAAAAGA	2486
Db	1382	TCTGATTGGAGCTGGACACAGCAGTGAGTTCAAAACCCATGACACAGAAATGAACAAAAGA	1441
Qy	2487	AGTCATTAGTAAATTTTCGACTGGAAAATCAATCTGCTTATCGCTACACAGTGGCAGA	2546
Db	1442	AGTCATTAGTAAATTTTCGACTGGAAAATCAATCTGCTTATCGCTACACAGTGGCAGA	1500
Qy	2547	AGAAGGTCTGGATTTAAAGAAATGTAACATTTGTTATCCGTTATGGTCTCGTCACCAATGA	2606
Db	1501	AGAAGGTCTGGATTTAAAGAAATGTAACATTTGTTATCCGTTATGGTCTCGTCACCAATGA	1560
Qy	2607	AATAGCCATGCTCCAGGCCGCTGAGAGCTGAGAGCTGATGAGAGCACCTACGTCCTGCT	2666
Db	1561	AATA -CCATGGTCCA -GCCCGTGGTTCGAGCTAGAGCTGATGAGAGCACCTACGCT -CTGCT	1617
Qy	2667	TGCTCACAGTGGTTTCAGGAGTTATCGAACATGAGACAGTTAATGATTTCCGAGAGAAAGAT	2726
Db	1618	TGCTCACAGTGGTTTCAGGAGTTATCGAACATGAGACAGTTAATGATTTCCGAGAGAAAGAT	1677
Qy	2727	GATGTATAAGCTATACATTTGTTCAAAATATGAAACCCAGAGGAGTATGCTCATAAAGAT	2786
Db	1678	GATGTATAAGCTATACATTTGTTCAAAATATGAAACCCAGAGGAGTATGCTCATAAAGAT	1734
Qy	2787	TTTGGAATTCAGATGCAAAAGTATAATGGAAAAGAAAATGAAAACCAAGAGAAATATTGC	2846
Db	1735	TTTGGAATTCAGATGCAAAAGTATAATGGAAAAGAAAATGAAAACCAAGAGAAATATTGC	1794
Qy	2847	CAAGCATTTACAAGAAATAACCCATCACTAATAACTTTTCCCTTTGCAAAAACCTGCAGTGTCT	2906
Db	1795	CAAGCATTTACAAGAAATAACCCATCACTAATAACTTTTCCCTTTGCAAAAACCTGCAGTGTCT	1854
Qy	2907	AGCCTGTTCTGGGAAGATATCCATGTAATTGAGAAAATGCATCAGTCAATATGACCCC	2966
Db	1855	AGCCTGTTCTGGGAAGATATCCATGTAATTGAGAAAATGCATCAGTCAATATGACCCC	1914
Qy	2967	AGAATTCAGGAACCTTTACATTGTAAGAGAAAAACAAAGCACTGCAAAAAGAA	3017
Db	1915	AGAATTCAGGAACCTTTACATTGTAAGAGAAAAACAAAGCACTGCAAAAAGAA	1965

RESULT 10
ADA53260
ID ADA53260 standard; cDNA; 1776 BP.
XX
AC ADA53260;
XX
DT 20-NOV-2003 (first entry)

XX	Human coding sequence, SEQ ID 828.	
DE		
XX	Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;	
KW	Gene Therapy; human; secretory protein; membrane proteins; cancer;	
KW	Inflammatory disease; osteoporosis; neurological disease; gene; ss.	
XX		
OS	Homo sapiens.	
XX		
PN	EP1293569-A2.	
XX		
PD	19-MAR-2003.	
XX		
PF	21-MAR-2002; 2002EP-00006586.	
XX		
PR	14-SEP-2001; 2001JP-00328381.	
PR	24-JAN-2002; 2002US-0350435P.	
XX		
XX	(HELI-) HELIX RES INST.	
PA	(REAS-) RES ASSOC BIOTECHNOLOGY.	
PA		
XX	Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;	
PI	Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;	
PI	Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;	
XX		
XX	WPI; 2003-395539/38.	
DR	P-PSDB; ADA54899.	
XX		
XX	New polynucleotides encoding full-length polypeptides, e.g. secretory	
PT	and/or membrane proteins, useful for developing medicines for diseases in	
PT	which the gene is involved, or as target molecules for gene therapy.	
XX		
PS	Claim 1; SEQ ID NO 828; 205pp; English.	
XX		
CC	The present invention relates to novel human secretory or membrane	
CC	proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-	
CC	ADA54071). The coding sequences are useful in the gene therapy of	
CC	diseases caused by abnormalities of the proteins, e.g. cancer,	
CC	inflammatory diseases, osteoporosis or neurological disease.	
XX		
XX	Sequence 1776 BP; 666 A; 291 C; 363 G; 456 T; 0 U; 0 Other;	
SQ		
Query Match 52.5%; Score 1766; DB 10; Length 1776;		
Best Local Similarity 99.7%; Pred. No. 0;		
Matches 1769; Conservative 0; Mismatches 5; Indels 0; Gaps 0;		
QY	1592 ACAACCCAGTGCATCCCTTCCCTCAGATACTGGGACTAACAGCTTCACCTGGTGGAG 1651	
Db	2 ACAACCCAGTGCATCCCTTCCCTCAGATACTGGGACTAACAGCTTCACCTGGTGGAG 61	
QY	1652 GGGCCACGAAGCAAGCCAAAGCTGAAGAACACATTTAAACCTATGTGCCAATCTTGATG 1711	
Db	62 GGGCCACGAAGCAAGCCAAAGCTGAAGAACACATTTAAACCTATGTGCCAATCTTGATG 121	
QY	1712 CATTACTATTAAACTGTTAAAGAAACCTTGATCACTGAAAAACCAATACAGGAGC 1771	
Db	122 CATTACTATTAAACTGTTAAAGAAACCTTGATCACTGAAAAACCAATACAGGAGC 181	
QY	1772 CATGCAAGAAGTTTGCCATTGCAGATGCAACCAAGAGATCCATTAAAGAAACCTTC 1831	
Db	182 CATGCAAGAAGTTTGCCATTGCAGATGCAACCAAGAGATCCATTAAAGAAACCTTC 241	
QY	1832 TAGAATAATGACAAGGATTCAAACTTATTGTCAAATGAGTCCAATGTGAGATTTGGAA 1891	
Db	242 TAGAATAATGACAAGGATTCAAACTTATTGTCAAATGAGTCCAATGTGAGATTTGGAA 301	
QY	1892 CTCAACCCATGAACAATGGGCCATTCAAAATGGAAAAAAGCTGCAAAAAAAGGAAATC 1951	
Db	302 CTCAACCCATGAACAATGGGCCATTCAAAATGGAAAAAAGCTGCAAAAAAAGGAAATC 361	
QY	1952 GCAAGAACCGTGTGTCGAGAACATTTGAGGAAGTACAATGAGGCCCTTACAAATTAATG 2011	
Db	362 GCAAGAACCGTGTGTCGAGAACATTTGAGGAAGTACAATGAGGCCCTTACAAATTAATG 421	

QY	2012	ACACAATTCGAATGATAGATCGGTATACCTCATCTTGAAACTTTCTATAATGAAGAGAAAG	2071
Db	422	ACACAATTCGAATGATAGATCGGTATACCTCATCTTGAAACTTTCTATAATGAAGAGAAAG	481
QY	2072	ATAAGAAGTTTGCAGTCATAGAAGATGATAGTATGAGGGTGGTGGTATGATGATGATTTGTG	2131
Db	482	ATAAGAAGTTTGCAGTCATAGAAGATGATAGTATGAGGGTGGTGGTATGATGATGATTTGTG	541
QY	2132	ATGGTGAATGAAGATGAGGATGATTTAAAGAAACCTTTGAAACTGGATGAAACAGATAGAT	2191
Db	601	ATGGTGAATGAAGATGAGGATGATTTAAAGAAACCTTTGAAACTGGATGAAACAGATAGAT	601
QY	2192	TTCTCATGACTTTTATTTTGGAAAAACAATAAAATGTTGAAAAAGCTGGCTGAAAAACCCAG	2251
Db	661	TTCTCATGACTTTTATTTTGGAAAAACAATAAAATGTTGAAAAAGCTGGCTGAAAAACCCAG	661
QY	2252	AATATGAAAAATGAAAAAGCTGACCAAAATTAAGAAATACCAATATGGAGCAATATACTAGGA	2311
Db	721	AATATGAAAAATGAAAAAGCTGACCAAAATTAAGAAATACCAATATGGAGCAATATACTAGGA	721
QY	2312	CTGAGGAATCAGCACGAGGAATAATCTTTACAAAAACACGACAGAGTGCATATGCGCTTT	2371
Db	781	CTGAGGAATCAGCACGAGGAATAATCTTTACAAAAACACGACAGAGTGCATATGCGCTTT	781
QY	2372	CCCAGTGGATTACTGAAAAATGAAAAATTTGCTGAAGTAGGAGTCAAAGCCACCATCTGA	2431
Db	841	CCCAGTGGATTACTGAAAAATGAAAAATTTGCTGAAGTAGGAGTCAAAGCCACCATCTGA	841
QY	2432	TTGGAGCTGGACACAGCAGTGAAGTTCAAACCCATGACACAGAAATGAACAAAAAGAGTCA	2491
Db	901	TTGGAGCTGGACACAGCAGTGAAGTTCAAACCCATGACACAGAAATGAACAAAAAGAGTCA	901
QY	2492	TTAGTAAATTTGCGCACTGGAAAAATCAATCTGCTTATCGTACCACAGTGGCAGAGAAG	2551
Db	961	TTAGTAAATTTGCGCACTGGAAAAATCAATCTGCTTATCGTACCACAGTGGCAGAGAAG	961
QY	2552	GTCTGGATATTAAAGAATGTAACATTTGTTATCCGTTATGGTCTCGTCACCAATGAATAG	2611
Db	1021	GTCTGGATATTAAAGAATGTAACATTTGTTATCCGTTATGGTCTCGTCACCAATGAATAG	1021
QY	2612	CCATGGTCCAGGCCCGTGGTCGAGCCAGAGCTGATGAGAGCACCTACGTCCTGGTTGCTC	2671
Db	1081	CCATGGTCCAGGCCCGTGGTCGAGCCAGAGCTGATGAGAGCACCTACGTCCTGGTTGCTC	1081
QY	2672	ACAGTGGTTCAGGAGTTATCGAACATGAGACAGTTAAATGATTTCCGAGAGAGATGATGT	2731
Db	1141	ACAGTGGTTCAGGAGTTATCGAACATGAGACAGTTAAATGATTTCCGAGAGAGATGATGT	1141
QY	2732	ATAAGCTATACATTGTGTTCAAAATATGAAACCAAGAGGAGTATGCTCATAGATTTTGG	2791
Db	1201	ATAAGCTATACATTGTGTTCAAAATATGAAACCAAGAGGAGTATGCTCATAGATTTTGG	1201
QY	2792	AATTACAGATGCAAAAGTATAATGGAAAAAGAAATGAAAAACCAAGAGAAATATTGCCAAGC	2851
Db	1261	AATTACAGATGCAAAAGTATAATGGAAAAAGAAATGAAAAACCAAGAGAAATATTGCCAAGC	1261
QY	2852	ATTACAAGAATAACCCATCACTAATAACTTTCCCTTTGCAAAAACTGCAGTGTGCTAGCCT	2911
Db	1321	ATTACAAGAATAACCCATCACTAATAACTTTCCCTTTGCAAAAACTGCAGTGTGCTAGCCT	1321
QY	2912	GTTCTGGGGAAGATATCCATGTAATTGAGAAAAATGCATCACGTCAATATGACCCAGAAT	2971
Db	1381	GTTCTGGGGAAGATATCCATGTAATTGAGAAAAATGCATCACGTCAATATGACCCAGAAT	1381
QY	2972	TCAAGGAACCTTTACATTGTAAGAGAAAAACAAAGCACTGCAAAAAGAGTGTGCCGACTATC	3031
Db	1441	TCAAGGAACCTTTACATTGTAAGAGAAAAACAAAGCACTGCAAAAAGAGTGTGCCGACTATC	1441
QY	3032	AAATAAATGGTGAAATCATCTGCAAAATGTGGCCAGGCTTGGGGAAACAATGATGGTGCACA	3091
Db	1501	AAATAAATGGTGAAATCATCTGCAAAATGTGGCCAGGCTTGGGGAAACAATGATGGTGCACA	1501

QY 3092 AAGGCTTAGATTGGCTTGCTCTCAAAATAAGGAATTTTGTAGTGGTTTCAAAAATAAAT 3151
Db |||||||
QY 1502 AAGGCTTAGATTGGCTTGCTCTCAAAATAAGGAATTTTGTAGTGGTTTCAAAAATAAAT 1561
Db |||||||
QY 3152 CAACAAAGAAACAATACAAAAAGTGGGTAGAAATACCTATCACATTTCCCAATCTTGACT 3211
Db |||||||
QY 1562 CAACAAAGAAACAATACAAAAAGTGGGTAGAAATACCTATCACATTTCCCAATCTTGACT 1621
Db |||||||
QY 3212 ATTCAGAAATGCTGTTTATTTAGTGATGAGGATTAGCACTTGATTGAAGATTCTTTTAAAA 3271
Db |||||||
QY 1622 ATTCAGAAATGCTGTTTATTTAGTGATGAGGATTAGCACTTGATTGAAGATTCTTTTAAAA 1681
Db |||||||
QY 3272 TACTATCAGTTAAACATTTAATATGATTATGATTAAATGATTATTCATTATGCTACAGAACTG 3331
Db |||||||
QY 1682 TACTATCAGTTAAACATTTAATATGATTATGATTAAATGATTATTCATTATGCTACAGAACTG 1741
Db |||||||
QY 3332 ACATAAGAATCAATAAAATGATTGTTTACTCTG 3365
Db |||||||
QY 1742 ACATAAGAATCAATAAAATGATTGTTTACTCTG 1775
Db |||||||
RESULT 11
AAI59285
ID AAI59285 standard; cDNA; 1557 BP.
XX
AC AAI59285;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 1488.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; Chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US034263.
XX
PR 23-DEC-1999; 99US-00471275.
PR 21-JAN-2000; 2000US-00488725.
PR 25-APR-2000; 2000US-00552317.
PR 20-JUN-2000; 2000US-00598042.
PR 19-JUL-2000; 2000US-00620312.
PR 03-AUG-2000; 2000US-00653450.
PR 14-SEP-2000; 2000US-00662191.
PR 19-OCT-2000; 2000US-00693036.
PR 29-NOV-2000; 2000US-00727344.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
PI Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI; 2001-442253/47.
DR P-PSDB; AAM40129.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders such
PT as central nervous system injuries.
XX
PS Claim 1; SEQ ID NO 1488; 10078pp; English.
XX
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and the
CC encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful

CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification
XX
SQ Sequence 1557 BP; 581 A; 248 C; 320 G; 408 T; 0 U; 0 Other;
Query Match 43.4%; Score 1460.2; DB 4; Length 1557;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1476; Conservative 0; Mismatches 8; Indels 1; Gaps 1;
QY 1882 GATTTTGGAACTCAACCCCTATGAACAATGGGCCATTCAATGGAATAAAAGCTGCAAAA 1941
Db |||||||
QY 1942 AAAGGAAATCGCAAAAGAACGTGTTTGTGCAGAACATTTTGAGGAAGTACAATGAGGCCCTA 2001
Db |||||||
QY 121 GAAGGAAATCGCGAGAACGTGTTTGTGCAGAACATTTTGAGGAAGTACAATGAGGCCCTA 180
QY 2002 CAAATTAATGACACAATTCGAATGATAGATGCGTATACTCATCTTTGAAAACCTTTCTATAAT 2061
Db |||||||
QY 181 CAAATTAATGACACAATTCGAATGATAGATGCGTATACTCATCTTTGAAAACCTTTCTATAAT 240
QY 2062 GAAGAGAAAGATAAGAGTTTTCAGTCATAGAAGATGATAGTGGTGGTGATGAT 2121
Db |||||||
QY 241 GAAGAGAAAGATAAGAGTTTTCAGTCATAGAAGATGATAGTGGTGGTGATGAT 300
QY 2122 GAGTATTGTGATGGTGATGAAGATGAGCATGATTTTAAAGAAACCTTTGAAAACCTGGATGAA 2181
Db |||||||
QY 301 GAGTATTGTGATGGTGATGAACATGAGGATGATTTTAAAGAAACCTTTGAAAACCTGGATGAA 360
QY 2182 ACAGATA-GATTTCTCATGACTTTTATTTTGTGAAAACAATAAATGTTGAAAAGGCTGGC 2240
Db |||||||
QY 361 ACAGATACGATTTCTCATGACTTTTATTTTGTGAAAACAATAAAGTTGAAAAGGCTGGC 420
QY 2241 TGAACACCCAGAAATATGAAAAATGAAAAGCTGACCAAAATTAAGAAATACCATAATGGAGCA 2300
Db |||||||
QY 421 TGAACACCCAGAAATATGAAAAATGAAAAGCTGACCAAAATTAAGAAATACCATAATGGAGCA 480
QY 2301 ATATACTAGGACTGAGGAATCAGCAGGAGGAATAATCTTTACAAAAACACGACAGAGTGC 2360
Db |||||||
QY 481 ATATACTAGGACTGAGGAATCAGCAGGAGGAATAATCTTTACAAAAACACGACAGAGTGC 540
QY 2361 ATATGCGCTTTCCAGTGGATTACTGAAAAATGAAAAATTTTGTGAAGTAGGAGTCAAAGC 2420
Db |||||||
QY 541 ATATGCGCTTTCCAGTGGATTACTGAAAAATGAAAAATTTTGTGAAGTAGGAGTCAAAGC 600
QY 2421 CCACCATCTGATTGGAGCTGGACACAGCAGTGAGTTCAAACCCCATGACACAGAATGAACA 2480
Db |||||||
QY 601 CCACCATCTGATTGGAGCTGGACACAGCAGTGAGTTCAAACCCCATGACACAGAATGAACA 660
QY 2481 AAAAGAAGTCAATTAGTAAATTTTCGCACTGGAAAAATCAATCTGCTTATCGTACCACAGT 2540
Db |||||||
QY 661 AAAAGAAGTCAATTAGTAAATTTTCGCACTGGAAAAATTAATCTGCTTATCGTACCACAGT 720
QY 2541 GGCAGAAGAGTCTGGATATTAAAGAATGTAACATTTGTTATCCGTTATGGTCTCGTCAC 2600
Db |||||||
QY 721 GGCAGAAGAGTCTGGATATTAAAGAATGTAACATTTGTTATCCGTTATGGTCTCGTCAC 780
QY 2601 CAATGAAATAGCCATGGTCCAGGCCCGTGGTCGAGCCAGAGCTGATGAGAGCACCTACGT 2660
Db |||||||
QY 781 CAATGAAATAGCCATGGTCCAGGCCCGTGGTCGAGCCAGAGCTGATGAGAGCACCTACGT 840
QY 2661 CCTGGTTGCTCACAGTGGTTTCAGGAGTTATCGAACATGAGACAGTTAATGATTTCCGAGA 2720
Db |||||||

Db 841 CCTGGTGTCTCACAGTGGTTCAGGAGTTATCGAACATGAGACAGTGAATGATTTCCGAGA 900

QY 2721 GAAGATGATGTATAAGCTATACATTTGTTCACAAATATGAACCAGAGGAGTATGCTCA 2780

Db 901 GAAGATGATGTATAAGCTATACATTTGTTCACAAATATGAACCAGAGGAGTATGCTCA 960

QY 2781 TAAGATTTTGGATTACAGATGCAAAAGTATAATGGAAAAAGAAATGAACCAAGAGAAA 2840

Db 961 TAAGATTTTGGATTACAGATGCAAAAGTATAATGGAAAAAGAAATGAACCAAGAGAAA 1020

QY 2841 TATTGCCAAGCATTACAAGAAATBAACCCATCACTAATAACTTTCCTTTTGCAAAACTGCAG 2900

Db 1021 TATTGCCAAGCATTACAAGAAATBAACCCATCACTAATAACTTTCCTTTTGCAAAACTGCAG 1080

QY 2901 TGTGCTAGCTGTCTCTGGGAAGATATCCATGTAATTGAGAAATGCAATCAGTCAATAT 2960

Db 1081 TGTGCTAGCTGTCTCTGGGAAGATATCCATGTAATTGAGAAATGCAATCAGTCAATAT 1140

QY 2961 GACCCAGAAATCAAGGAACCTTTACATTTGTAAGAGAAAAAAGCACTGCAAAAAGAGTG 3020

Db 1141 GACCCAGAAATCAAGGAACCTTTACATTTGTAAGAGAAAAAAGCACTGCAAAAAGAGTG 1200

QY 3021 TGCCGACTATCAAAATAAATGGTGAATCATCTGCAAAATGTGGCCAGGCTTGGGAAACAAT 3080

Db 1201 TGCCGACTATCAAAATAAATGGTGAATCATCTGCAAAATGTGGCCAGGCTTGGGAAACAAT 1260

QY 3081 GATGGTGCAAAAGGCTTAGATTTGCCTTGCTCTCAAAATAAGGAATTTTGTAGTGGTTTT 3140

Db 1261 GATGGTGCAAAAGGCTTAGATTTGCCTTGCTCTCAAAATAAGGAATTTTGTAGTGGTTTT 1320

QY 3141 CAAAAATAATTCAACAAAGAAACAATACAAAAAGTGGGTAGAATTACCTATCACATTTCC 3200

Db 1321 CAAAAATAATTCAACAAAGAAACAATACAAAAAGTGGGTAGAATTACCTATCACATTTCC 1380

QY 3201 CAATCTTGACTATTACAGAAATGCTGTTTATTTAGTGATGAGGATAGCACTTGATTGAAGA 3260

Db 1381 CAATCTTGACTATTACAGAAATGCTGTTTATTTAGTGATGAGGATAGCACTTGATTGAAGA 1440

QY 3261 TTCTTTTAAATACTATCAGTTAAACATTTAATATGATTATGATTATGATTATGATTATG 3320

Db 1441 TTCTTTTAAATACTATCAGTTAAACATTTAATATGATTATGATTATGATTATGATTATG 1500

QY 3321 CTACAGAACTGACATAAGAAATCAATAAATGATTTTACTCTG 3365

Db 1501 CTACAGAACTGACATAAGAAATCAATAAATGATTTTACTCTG 1545

RESULT 12

ABA04916

ID ABA04916 standard; DNA; 1443 BP.

XX ABA04916;

DT 05-MAR-2002 (first entry)

DE Human RNA helicase RH116 related DNA sequence #3.

XX Human; RH116; RNA helicase; cytostatic; virucide; anti-HIV;

KW immunosuppressive; immunostimulatory; antirheumatic; antiarthritic;

KW antiarteriosclerotic; osteopathic; antidiabetic; hepatotropic;

KW antiinflammatory; cancer; infection; HIV; hepatitis; genetic disease;

KW autoimmune disease; graft rejection; vaccine; ds.

OS Homo sapiens.

XX WO200185955-A1.

PN 15-NOV-2001.

XX 11-MAY-2001; 2001WO-FR001441.

PF 11-MAY-2000; 2000FR-00006030.

PR

XX

PA (ISTA-) ISTAC.

PA (INSP) INST PASTEUR LILLE.

XX Bahr G, Cocude C, Capron A;

XX WPI; 2002-082898/11.

XX New polypeptide, useful for treating and diagnosing cancer or

PT inflammation, and drug screening, comprises a human polynucleotide

PT homologous to RNA helicase.

XX Example 2; Page 94-95; 114pp; French.

PS The present invention relates to human RH116 (see AAM47798). RH116 is a

XX 116kDa protein and has homology to RNA helicases (DEXH box). RH116 and

CC its coding sequence are useful for treating cancer; acute or chronic

CC infections (especially by HIV or hepatitis B or C); inherited genetic

CC diseases; (auto)immune diseases (particularly rheumatism, arthritis,

CC arteriosclerosis, osteoporosis and diabetes, but many others listed), and

CC to prevent graft rejection. RH116 and its coding sequence are also useful

CC for inducing, or increasing, the immune response to a vaccine. The

CC present sequence was used in an example from the present invention

XX Sequence 1443 BP; 508 A; 261 C; 313 G; 360 T; 0 U; 1 Other;

SQ

Query Match 42.7%; Score 1435.6; DB 6; Length 1443;

Best Local Similarity 99.7%; Pred. No. 0;

Matches 1438; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 687 GAAAGAAAACCTGGTTCTCTGCAATTTCTGAATGTTCTTCGTAACAGGAAACAATGAAC 746

Db 1 GAAAGAAAACCTGGTTCTCTGCAATTTCTGAATGTTCTTCGTAACAGGAAACAATGAAC 60

QY 747 TGTCCAAGAGTTAAACAGGCTCTGATTGCTCAGAAAAGCAATGAGAGATTGAGAAATTC 806

Db 61 TGTCCAAGAGTTAAACAGGCTCTGATTGCTCAGAAAAGCAATGAGAGATTGAGAAATTC 120

QY 807 ACAAGTTGATGGTCTCAAGTGAAGAGCAAACTTCTTCAACACACAGTTTCAGCCAAATCT 866

Db 121 ACAAGTTGATGGTCTCAAGTGAAGAGCAAACTTCTTCAACACACAGTTTCAGCCAAATCT 180

QY 867 GGAGAAGGAGGCTCTGGGCGATGGAGAATAACTCATCAGAATCATCTTTGCAGATTTCTTC 926

Db 181 GGAGAAGGAGGCTCTGGGCGATGGAGAATAACTCATCAGAATCATCTTTGCAGATTTCTTC 240

QY 927 TGTAGTTTCAGAAATCAGACACAAGTTTGGCAGAAAGAGTGTCTGCTTAGATGAAAG 986

Db 241 TGTAGTTTCAGAAATCAGACACAAGTTTGGCAGAAAGAGTGTCTGCTTAGATGAAAG 300

QY 987 TCTTGGACATAACAGCAACATGGGCGATGATTCAGGCCACCATGGGAAGTATTGAGATGA 1046

Db 301 TCTTGGACATAACAGCAACATGGGCGATGATTCAGGCCACCATGGGAAGTATTGAGATGA 360

QY 1047 AGAGAATGTGGCAGCAAGAGCATCCCGGAGCCAGAACTCCAGCTCAGGCCCTTACCAAAT 1106

Db 361 AGAGAATGTGGCAGCAAGAGCATCCCGGAGCCAGAACTCCAGCTCAGGCCCTTACCAAAT 420

QY 1107 GGAAGTTGCCAGCCAGCCTTGAAGGGGAAGAAATATCATCATCTGCCCTCCCTACAGGGAG 1166

Db 421 GGAAGTTGCCAGCCAGCCTTGAAGGGGAAGAAATATCATCATCTGCCCTCCCTACAGGGAG 480

QY 1167 TGGAAAAACCCAGAGTGGCTGTTTACATTTGCCAAGGATCACTTAGACAGAAAGAAAAGC 1226

Db 481 TGGAAAAACCCAGAGTGGCTGTTTACATTTGCCAAGGATCACTTAGACAGAAAGAAAAGC 540

QY 1227 ATCTGAGCCTGGAAAAAGTTATAGTTCTTGTCAATAAGGTACTGCTAGTTGAACAGCTCTT 1286

Db 541 ATCTGAGCCTGGAAAAAGTTATAGTTCTTGTCAATAAGGTACTGCTAGTTGAACAGCTCTT 600

QY 1287 CCGCAAGGAGTTCCAAACCACTTTTGAAGAAATGGTATCGTGTATTGGATTAAGTGGTGA 1346

Db 601 CCGCAAGGAGTTCCAAACCACTTTTGAAGAAATGGTATCGTGTATTGGATTAAGTGGTGA 660

Db 253 CTGAACACCCCTAATATGAAATGAAAAGCTGACCAAATTAAGAAATACCATATATGGAGC 312

QY 2300 AATATACTAGGACTGAGGAATCAGCACGAGGAATATCTTTACAAAAACACGACAGAGTG 2359

Db 313 AATATACTAGGACTGAGGAATCAGCACGAGGAATATCTTTACAAAAACACGACAGAGTG 372

QY 2360 CATATGCGCTTTCCAGTGGATTACTGAAAATGAAAATTTGCTGAAGTAGGAGTCAAAG 2419

Db 373 CATATGCGCTTTCCAGTGGATTACTGAAAATGAAAATTTGCTGAAGTAGGAGTCAAAG 432

QY 2420 CCCACCATCTGATTGGAGCTGGACACAGCAGTGAGTTCAAACCCCATGACACAGAATGAAC 2479

Db 433 CCCACCATCTGATTGGAGCTGGACACAGCAGTGAGTTCAAACCCCATGACACAGAATGAAC 492

QY 2480 AAAAGAAGTCAATTAGTAAATTTGCACTGGAAAAATCAATCTGCTTATCGTACCACAG 2539

Db 493 AAAAGAAGTCAATTAGTAAATTTGCACTGGAAAAATCAATCTGCTTATCGTACCACAG 552

QY 2540 TGGCAGAAGAAGTCTGGATATTAAAGAATGTAACATTTATCCGTTATGGTCTCGTCA 2599

Db 553 TGGCAGAAGAAGTCTGGATATTAAAGAATGTAACATTTATCCGTTATGGTCTCGTCA 612

QY 2600 CCAATGAAATAGCCATGGTCCAGGCCCGTGGTCGAGCCAGAGCTGATGAGACACCTACG 2659

Db 613 CCAATGAAATAGCCATGGTCCAGGCCCGTGGTCGAGCCAGAGCTGATGAGACACCTACG 672

QY 2660 TCCTGGTTGCTCACAGTGGTTCCAGGAGTTATCGAACATGAGACAGTTAATGATTTCCGAG 2719

Db 673 TCCTGGTTGCTCACAGTGGTTCCAGGAGTTATCGAACATGAGACAGTTAATGATTTCCGAG 732

QY 2720 AGAAGATGATGTATAAAGCTATACATTTGTGTTCAAATATGAACACAGAGGAGTATGCTC 2779

Db 733 AGAAGATGATGTATAAAGCTATACATTTGTGTTCAAATATGAACACAGAGGAGTATGCTC 792

QY 2780 ATAAGATTTTGGAAATTACAGATGCAAAAGTATAATGGAAAAGAAAATGAAACCAAGAGAA 2839

Db 793 ATAAGATTTTGGAAATTACAGATGCAAAAGTATAATGGAAAAGAAAATGAAACCAAGAGAA 852

QY 2840 ATATTGCCAAGCATTAACAAGATAACCCATCACTAATAACTTTCTTTGCAAAAACCTGCA 2899

Db 853 ATATTGCCAAGCATTAACAAGATAACCCATCACTAATAACTTTCTTTGCAAAAACCTGCA 912

QY 2900 GTGTGCTAGCCTGTTCTGGGGAAGATATCCATGTAAATTGAGAAAATGCATCAGTCAATA 2959

Db 913 GTGTGCTAGCCTGTTCTGGGGAAGATATCCATGTAAATTGAGAAAATGCATCAGTCAATA 972

QY 2960 TGACCCGAGAAATCAAGGAATTTACATTGTGAAGAAAACAAGCACTGCAAAAAGAGT 3019

Db 973 TGACCCGAGAAATCAAGGAATTTACATTGTGAAGAAAACAAGCACTGCAAAAAGAGT 1032

QY 3020 GTCCGCACTATCAATAAATGTTGAAATCATCTGCAAAATGTGGCCAGCTTGGGGAACAA 3079

Db 1033 GTCCGCACTATCAATAAATGTTGAAATCATCTGCAAAATGTGGCCAGCTTGGGGAACAA 1092

QY 3080 TGATGGTGCAAAAGGCTTAGATTGCTTGTCTCAAAATAAGGAATTTGTAGTGGTTT 3139

Db 1093 TGATGGTGCAAAAGGCTTAGATTGCTTGTCTCAAAATAAGGAATTTGTAGTGGTTT 1152

QY 3140 TCAAAAATAATTCAACAAGAAACAATAACAAAAGTGGGTAGAAATTACCTATCACATTC 3199

Db 1153 TCAAAAATAATTCAACAAGAAACAATAACAAAAGTGGGTAGAAATTACCTATCACATTC 1212

QY 3200 CCAATCTTGACTATTGAGAAATGCTGTTTATTATTAGTATGAGGATTAGCACTTGATTGAAG 3259

Db 1213 CCAATCTTGACTATTGAGAAATGCTGTTTATTATTAGTATGAGGATTAGCACTTGATTGAAG 1272

QY 3260 ATTCTTTTAAATACTATCAGTTAAACATTTAATATGATTATGATTATGATT-CATTA 3318

Db 1273 ATTCTTTTAAATACTATCAGTTAAACATTTAATATGATTATGATTATGATTATGATTCCATAA 1332

QY 3319 TGCTACAGAACTGACATAAGAAATCAATAAAATGATTGTTTACTCTG 3365

Db 1333 TGCTACAGAACTGACATAAGAAATCAATAAAATGATTGTTTACTCTG 1379

RESULT 14
ADC32572

ID ADC32572 standard; cDNA; 1382 BP.

XX ADC32572;

XX 18-DEC-2003 (first entry)

XX Human novel cDNA contig sequence, SEQ ID NO:2654.

DE Human; diagnostic; drug screening; forensics; gene mapping;
XX biodiversity assessment; Parkinson's disease; Alzheimer's disease;
KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
KW ulcers; osteoporosis; autoimmune disease; cancer;
KW molecular weight marker; food supplement; antiparkinsonian; nootropic;
KW neuroprotective; antianaemic; anticoagulant; thrombolytic; vulnerary;
KW antiulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
KW gene therapy; ss.

XX Homo sapiens.

XX WO2003029271-A2.

XX 10-APR-2003.

XX 24-SEP-2002; 2002WO-US030474.

XX 24-SEP-2001; 2001US-0324631P.

XX (HYSE-) HYSEQ INC.

XX Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;
PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
PI Haley-Vicente D, Drmanac RT;

XX WPI; 2003-371981/35.

XX P-PSDB; ADC33339.

XX New polynucleotide and polypeptide useful for diagnosing, preventing or
PT treating conditions such as neurodegenerative diseases, anemias, platelet
PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
PT cancer.

XX Example 2; SEQ ID NO 2654; 1185pp; English.

XX The invention relates to 971 novel human cDNA sequences (ADC29919-
CC ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The
CC invention also relates to nucleic acid sequences over 99% identical with
CC the novel human cDNAs. The invention additionally encompasses expression
CC vectors and host cells comprising a nucleic acid of the invention; the
CC recombinant production of a polypeptide of the invention; an antibody
CC against a polypeptide of the invention; a method of detecting of
CC polynucleotides or polypeptides of the invention; and methods of
CC identifying a compound which binds to a polypeptide of the invention. The
CC invention further discloses methods of preventing, treating or
CC ameliorating a medical condition; kits comprising polynucleotide probes
CC and/or monoclonal antibodies for carrying out the methods of the
CC invention; methods for the identification of compounds that modulate the
CC expression or activity of the polynucleotide and/or polypeptide; and 767
CC contig sequences corresponding to the cDNA sequences of the invention
CC (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628
CC -ADC33394). The nucleic acids and polypeptides of the invention are
CC useful in diagnostics, drug screening, forensics, gene mapping, in the
CC identification of mutations responsible for genetic disorders or other
CC traits, for assessing biodiversity, and in producing many other types of
CC data and products dependent on DNA and amino acid sequences. They are
CC also used for treating diseases such as Parkinson's disease, Alzheimer's
CC disease and other neurodegenerative diseases, anaemia, platelet
CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
CC cancer. The nucleic acids may also be used as hybridisation probes or
CC primers, and in the recombinant production of a protein. The polypeptides

CC are also useful in generating antibodies, as molecular weight markers,
CC and as food supplements. The present sequence represents a human contig
CC sequence used in an example of the invention. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX

SQ Sequence 1382 BP; 513 A; 217 C; 285 G; 367 T; 0 U; 0 Other;

Query Match 39.1%; Score 1316.6; DB 10; Length 1382;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1355; Conservative 0; Mismatches 9; Indels 3; Gaps 3;

QY 2002 CAAATTAATGACACAATTCGAATGATAGATGCGTATACCTCATCTTGAAAACCTTTCTATAAT 2061
Db CAAATTAATGACACAATTCGAATGATAGATGCGTATACCTCATCTTGAAAACCTTTCTATAAT 72
QY 2062 GAAGAGAAAGATAAGAAAGTTTGCAGTCATAGAAGATGATAGTGGGTGGTGATGAT 2121
Db GAAGAGAAAGATAAGAAAGTTTGCAGTCATAGAAGATGATAGTGGGTGGTGATGAT 132
QY 2122 GAGTATTGTGATGGTGAAGATGAGGATGATTTAAAGAAACCTTTGAAACTGGATGAA 2181
Db GAGTATTGTGATGGTGAACATGAGGATGATTTTAAAGAAACCTTTGAAACTGGATGAA 192
QY 2182 ACAGATA-GATTTCTCATGACTTTTATTTTGAAAAACAATAAAT-GTTGAAAAGGCTGG 2239
Db ACAGATACGATTTCTCATGACTTTTATTTTGAACACAATAAATAGTTGAATAGGCTGG 252
QY 2240 CTGAAAAACCCAGAATATGAAAAATGAAAAGCTGACCAAAATTAAGAAATACCATAATGGAGC 2299
Db CTGAACACCCCTAATATGAAAATGAAAAGCTGACCAAAATTAAGAAATACCATAATGGAGC 312
QY 2300 AATATACTAGACTGAGGAATCAGCAGGAGGAATAATCTTTACAAAAACACGACAGAGTG 2359
Db AATATACTAGACTGAGGAATCAGCAGGAGGAATAATCTTTACAAAAACACGACAGAGTG 372
QY 2360 CATATGCGCTTTCCAGTGGATTACTGAAAAATGAAAAATTTGCTGAAGTAGGAGTCAAAAG 2419
Db CATATGCGCTTTCCAGTGGATTACTGAAAAATGAAAAATTTGCTGAAGTAGGAGTCAAAAG 432
QY 2420 CCCACCATCTGATTTGGAGCTGGACACAGCAGTGAGTTCAAAACCCATGACACAGAAATGAAC 2479
Db CCCACCATCTGATTTGGAGCTGGACACAGCAGTGAGTTCAAAACCCATGACACAGAAATGAAC 492
QY 2480 AAAAAGAAAGTCATTAGTAAATTTCCGACTGGAAAAATCAATCTGCTTATCGCTACCACAG 2539
Db AAAAAGAAAGTCATTAGTAAATTTCCGACTGGAAAAATCAATCTGCTTATCGCTACCACAG 552
QY 2540 TGGCAGAAAGAGGTCTGGATATTAAGAATGTAACATTTATCCGTTATGGTCTCGTCA 2599
Db TGGCAGAAAGAGGTCTGGATATTAAGAATGTAACATTTATCCGTTATGGTCTCGTCA 612
QY 2600 CCAATGAATAGCCATGGTCCAGGCCGTTGTCGAGCCAGAGCTGATGAGAGCACCTACG 2659
Db CCAATGAATAGCCATGGTCCAGGCCGTTGTCGAGCCAGAGCTGATGAGAGCACCTACG 672
QY 2660 TCCTGGTTGCTCACAGTGGTTTCAGGAGTTATCGAACATGACAGAGTTAATGATTCCGAG 2719
Db TCCTGGTTGCTCACAGTGGTTTCAGGAGTTATCGAACATGACAGAGTTAATGATTCCGAG 732
QY 2720 AGAAGATGATGTATAAAGCTATACATTGTGTTCAAAAATATGAAACCCAGAGGAGTATGCTC 2779
Db AGAAGATGATGTATAAAGCTATACATTGTGTTCAAAAATATGAAACCCAGAGGAGTATGCTC 792
QY 2780 ATAAGATTTTGAATTACAGATGCAAGTATAATGGAAGAAAGAAATGAAAAACCAAGAGAA 2839
Db ATAAGATTTTGAATTACAGATGCAAGTATAATGGAAGAAAGAAATGAAAAACCAAGAGAA 852
QY 2840 ATATTGCCAAGCATTACAAGAATAACCCATCACTAATAACTTTCCTTTGCAAAAACCTGCA 2899
Db ATATTGCCAAGCATTACAAGAATAACCCATCACTAATAACTTTCCTTTGCAAAAACCTGCA 912

QY 2900 GTGTGCTAGCCTGTTCTGGGGAGAGATATCCATGTAAATTGAGAAAAATGCATCACGTCAATA 2959
Db GTGTGCTAGCCTGTTCTGGGGAGAGATATCCATGTAAATTGAGAAAAATGCATCACGTCAATA 972
QY 2960 TGACCCCGAGAATTCGAAGAACTTTACATTTGTAAGAGAAAAACAAGCACTGCAAAAAGAGT 3019
Db TGACCCCGAGAATTCGAAGAACTTTACATTTGTAAGAGAAAAACAAGCACTGCAAAAAGAGT 1032
QY 3020 GTGCCGACTATCAAAATAAATGGTGAATCATCTGCAAAATGTGGCCAGGCTTGGGGAACAA 3079
Db GTGCCGACTATCAAAATAAATGGTGAATCATCTGCAAAATGTGGCCAGGCTTGGGGAACAA 1092
QY 3080 TGATGGTGCACAAAGGCTTAGATTGGCCTTGCTCTCAAAATAAGGAATTTTGTAGTGGTTT 3139
Db TGATGGTGCACAAAGGCTTAGATTGGCCTTGCTCTCAAAATAAGGAATTTTGTAGTGGTTT 1152
QY 3140 TCAAAAATAATTCAACAAAGAAACAATACAAAAAGTGGGTAGAAATTACCTATCACATTTT 3199
Db TCAAAAATAATTCAACAAAGAAACAATACAAAAAGTGGGTAGAAATTACCTATCACATTTT 1212
QY 3200 CCAATCTTGACTATTTCAGAAATGCTGTTTATTTAGTGATGAGGATTAGCACTTGATTGAAG 3259
Db CCAATCTTGACTATTTCAGAAATGCTGTTTATTTAGTGATGAGGATTAGCACTTGATTGAAG 1272
QY 3260 ATTCTTTTAAATACTACTATCAGTTTAAACATTTAATATGATGATTAATGTATT-CATTA 3318
Db ATTCTTTTAAATACTACTATCAGTTTAAACATTTAATATGATGATTAATGTATTCCATAA 1332
QY 3319 TGCTACAGAACTGACATAAAGAAATCAATAAAATGATTGTTTACTCTG 3365
Db TGCTACAGAACTGACATAAAGAAATCAATAAAATGATTGTTTACTCTG 1379

RESULT 15

ADI30722

ID ADI30722 standard; cDNA; 1392 BP.

XX AC ADI30722;

XX DT 17-JUN-2004 (first entry)

XX DE Human cDNA #48.

Human; gene; ss; immunological response; immunopathological condition;
Crohn's disease; asthma; ulcerative colitis; hypereosinophilia;
irritable bowel syndrome; osteoarthritis; rheumatoid arthritis;
acute monocytic leukaemia; antiinflammatory; antiasthmatic; antiulcer;
osteopathic; antiarthritic; antirheumatic; cytostatic.

OS Homo sapiens.

XX US6607879-B1.

XX PD 19-AUG-2003.

XX PF 09-FEB-1998; 98US-00023655.

XX PR 09-FEB-1998; 98US-00023655.

XX PA (INCY-) INCYTE CORP.

XX PI Cocks BG, Stuart SG, Seilhamer JJ;

XX DR WPI; 2003-895307/82.

XX A composition comprising a plurality of cDNAs, useful for detecting
PT altered expression of genes in an immunological response or for
PT diagnosing and treating an immunopathology, e.g. Crohn's disease, asthma
PT or osteoarthritis.

XX PS Claim 1; SEQ ID NO 48; 50pp; English.

XX CC The invention relates to a composition comprising a plurality of cDNAs

CC for detecting the altered expression of genes in an immunological
CC response. The invention also relates to a method of diagnosing or
CC monitoring the treatment of an immunopathological condition in a sample,
CC comprising obtaining nucleic acids from a sample, contacting the nucleic
CC acids of the sample with an array comprising the plurality of cDNAs under
CC conditions to form one or more hybridisation complexes, detecting the
CC hybridisation complexes and comparing the levels of the detected
CC hybridisation complexes with the level of hybridisation complexes
CC detected in a non-diseased sample, where an altered level of the detected
CC hybridisation complexes correlates with the presence of an
CC immunopathological condition. Also disclosed are an expression profile
CC comprising a microarray and a plurality of detectable complexes and a
CC method for identifying a plurality of polynucleotide probes. The cDNAs
CC are useful as hybridisable array elements in a microarray for monitoring
CC the expression of target polynucleotides. The microarray can be used in
CC the diagnosis of an immunopathology, such as Crohn's disease, asthma,
CC ulcerative colitis, hypereosinophilia, irritable bowel syndrome,
CC osteoarthritis, rheumatoid arthritis or acute monocytic leukaemia, and in
CC identifying agents for the treatment of the diseases. The microarray may
CC also be used in drug discovery and development, toxicological and
CC carcinogenicity studies, forensics or pharmacogenomics. The composition
CC may also be used in purification of a subpopulation of mRNAs, cDNAs or
CC genomic fragments. This sequence represents a human cDNA of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification but was obtained in electronic format directly
CC from USPTO at seqdata.uspto.gov/sequence.html.

XX	Sequence 1392 BP; 497 A; 215 C; 264 G; 351 T; 0 U; 65 Other;	
SQ		
	Query Match 38.0%; Score 1278.8; DB 11; Length 1392;	
	Best Local Similarity 94.1%; Pred. No. 1.2e-306;	
	Matches 1311; Conservative 0; Mismatches 77; Indels 5; Gaps 2;	
Qy	1976 ATTTGAGGAAGTACAAATGAGGCCCTACAAATTAATGACACAAATTCGAATGATAGATGCGT	2035
Db	1 ATTTGAGGAAGTACAAATGAGGCCCTACAAATTAATGACACAAATTCGAATGATAGATGCGT	60
Qy	2036 ATACTCATCTTGAAACTTCTATAATGAAGAGAAAGATAAGAGTTTGAGTCATAGAAG	2095
Db	61 ATACTCATCTTGAAACTTCTATAATGAAGAGAAAGATAAGAGTTTGAGTCATANNNN	120
Qy	2096 ATGATAGTGATGAGGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT	2155
Db	121 NNT	180
Qy	2156 TAAAGAAACCTTTGAAACTGATGAAACAGATAGATTTCTCATGACTTTATTTTGTGAAA	2215
Db	181 TAAAGAAACCTTTGAAACTGATGAAACAGATAGATTTCTCATGACTTTATTTTGTGAAA	240
Qy	2216 ACAATAAATGTTGAAAAGGCTGGCTGAAAACCCAGAAATATGAAAATGAAAAGCTGACCA	2275
Db	241 ACAATAAATGTTGAAAAGGCTGGCTGAAAACCCAGAAATATGAAAATGAAAAGCTGACCA	300
Qy	2276 AATTAAAGAAATACCATAATGGAGCAATATACCTAGGACTGAGGAATCAGCAGGAGGAATAA	2335
Db	301 AATTAAAGAAATACCATAATGGAGCAATATACCTAGGACTGAGGAATCAGCAGGAGGAATAA	360
Qy	2336 TCTTTACAAAACACGACAGAGTGATATGCGCTTTCCAGTGGATTACTGAAAATGAAA	2395
Db	361 TCTTTACAAAACACGACAGAGTGATATGCGCTTTCCAGTGGATTACTGAAAATGAAA	420
Qy	2396 AATTTGCTGAAGTAGGAGTCAAAGCCACCATCTGATTGGAGCTGGACACAGCAGTGAGT	2455
Db	421 AATTTGCTGAAGTAGGAGTCAAAGCCACCATCTGATTGGAGCTGGACACAGCAGTGAGT	480
Qy	2456 TCAAAACCCATGACACAGAATGAACAAAAGAAAGTCAATTAGTAAATTTCCGACTGGAAAAA	2515
Db	481 TCAAAACCCATGACACAGAATGAACAAAAGAAAGTCAATTAGTAAATTTCCGACTGGAAAAA	540
Qy	2516 TCAATCTGCTTATCGCTACCACAGTGGCAGAGAAAGGCTCTGGATATTTAAAGAAATGTAACA	2575
Db	541 TAAATCTGCTTATCGCTACCACAGTGGCAGAGAAAGGCTCTGGATATTTAAAGAAATGTAACA	600

Qy	2576 TTGTTATCCGTTATGGTCTCGTCCCAATGAAATAGCCATGGTCCAGGCCCGTGGTCGAG	2635
Db	601 TTGTTATCCGTTATGGTCTCGTCCCAATGAAATAGCCATGGTCCAGGCCCGTGGTCGAG	660
Qy	2636 CCAGAGCTGATGAGAGCACCTACGCTCTGGTTGCTCACAGTGGTTCAGGAGTTATCGAAC	2695
Db	661 CCAGAGCTGATGAGAGCACCTACGCTCTGGTTGCTCACAGTGGTTCAGGAGTTATCGAAC	720
Qy	2696 ATGAGACAGTTAATGATTTCCGAGAGAGATGATGTATAAAGCTATACATTGTGTTCAAA	2755
Db	721 GTGAGACAGTTAATGATTTCCGAGAGAGATGATGTATAAAGCTATACATTGTGTTCAAA	780
Qy	2756 ATATGAAACCCAGAGGAGTATGCTCATAAGATTTTGGAAATTACAGATGCAAAAGTATAATGG	2815
Db	781 ATATGAAACCCAGAGGAGTATGCTCATAAGATTTTGGAAATTACAGATGCAAAAGTATAATGG	840
Qy	2816 AAAAGAAATGAAACCCAGAGAAATATTGCCAAGCATTACAAGATAACCCATCACTAA	2875
Db	841 AAAAGAAATGAAACCCAGAGAAATATTGCCAAGCATTACAAGATAACCCATCACTAA	900
Qy	2876 TAACTTTCCTTTGCAAAAACCTGAGTGTGCTAGCTGTTCTGGGGAAGATATCCATGTAA	2935
Db	901 TAACTTTCCTTTGCAAAAACCTGAGTGTGCTAGCTGTTCTGGGGAAGATATCCATGTAA	960
Qy	2936 TTGAGAAATGCATCAGCTCAATATGACCCAGAAATTCAGGAACTTTACATTTGTAAGAG	2995
Db	961 TTGAGAAATGCATCAGCTCAATATGACCCAGAAATTCAGGAACTTTACATTTGTAAGAG	1020
Qy	2996 AAAACAAAGCACTGCAAAAGAGTGTGCCGACTATCAAAATAAATGGTGAATCATCTGCA	3055
Db	1021 AAAACAAAGCACTGCAAAAGAGTGTGCCGACTATCAAAATAAATGGTGAATCATCTGCA	1080
Qy	3056 AATGTGGCCAGGCTTGGGGAACAATGATGGTGACAAAGGCTTAGATTGGCCTTGTCTCA	3115
Db	1081 AATGTGGCCAGGCTTGGGGAACAATGATGGTGACAAAGGCTTAGATTGGCCTTGTCTCA	1140
Qy	3116 AAATAAGGAATTTGTAGTGGTTTCAAAAATAATTCACAAAGAAACAATAACAAAAAGT	3175
Db	1141 AAATAAGGAATTTGTAGTGGTTTCAAAAATAATTCACAAAGAAACAATAACAAAAAGT	1200
Qy	3176 GGGTAGAATTACCTATCACATTTCCCAATCTTGACTATTTCAGAAATGCTGTTTATTAGTG	3235
Db	1201 GGGTAGAATTACCTATCACATTTCCCAATCTTGACTATTTCAGAAATGCTGTTTATTAGTG	1260
Qy	3236 ATGAGGATTAGCACTTGATTGAAGATCTTTTAAATACIATCAGTTAAACATTATAATAT	3295
Db	1261 ATGAGGATTAGCACTTGATTGAAGATCTTTTAAATACIATCAGTTAAACATTATAATAT	1318
Qy	3296 GATTATGATTAATGTAATTCATTATGCTACAGAACTGACATAAGATC---AATAAATGA	3352
Db	1319 TGTTATGATTAATGTAATTCATTATGCTACAGAACTGACATAAATAATGA	1378
Qy	3353 TTGTTTACTCTG 3365	
Db	1379 TTGTTTACTCTG 1391	

Search completed: January 27, 2005, 11:31:03
Job time : 1449 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: January 27, 2005, 10:52:16 ; Search time 294 Seconds
(without alignments)
8135.390 Million cell updates/sec

Title: US-09-515-363C-1
Perfect score: 3365
Sequence: 1 gcgcgcggcctgagagccc.....aaaatgattgttttactctg 3365

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1278.8	38.0	1392	4	US-09-023-655-48
2	285.4	8.5	301	1	US-08-143-576-7
3	285.4	8.5	301	4	US-09-221-268D-12
4	179.6	5.3	2590	4	US-09-799-451-771
5	124.6	3.7	193	1	US-08-143-576-6
6	124.6	3.7	193	4	US-09-221-268D-11
C 7	69.4	2.1	7218	1	US-08-232-463-14
8	63	1.9	6184	4	US-09-590-968B-1
9	62	1.8	5852	4	US-09-853-768-10
10	62	1.8	7037	4	US-09-853-768-3
C 11	54.6	1.6	371	3	US-09-222-575-46
C 12	54.6	1.6	371	4	US-09-389-681-46
C 13	54.6	1.6	371	4	US-09-620-405B-46
C 14	54.6	1.6	371	4	US-09-339-338-46
C 15	54.6	1.6	371	4	US-09-433-826B-46
C 16	54.6	1.6	371	4	US-09-604-287A-46
C 17	54.6	1.6	371	4	US-09-285-480-46
C 18	54.6	1.6	371	4	US-09-834-759-46
C 19	54.6	1.6	371	4	US-09-590-751A-46
20	54.6	1.6	1141	4	US-09-806-708B-22
21	54.2	1.6	1964	4	US-10-140-002-403
22	53.2	1.6	1671	4	US-09-248-796A-4970
23	51.4	1.5	259	4	US-09-513-999C-2844
C 24	50.2	1.5	1141	4	US-09-806-708B-22
25	49.8	1.5	8920	2	US-08-446-855A-1
26	49.8	1.5	8920	3	US-09-150-741-1
27	48.2	1.4	2535	4	US-09-799-451-549

28	47.8	1.4	2319	3	US-09-058-489-90	Sequence 90, Appl
29	47.8	1.4	2451	4	US-09-976-594-786	Sequence 786, App
30	47.8	1.4	4416	3	US-09-058-489-17	Sequence 17, Appl
31	47.8	1.4	5361	3	US-08-973-462-2	Sequence 2, Appli
32	47.8	1.4	6152	3	US-08-973-462-1	Sequence 1, Appli
33	47.2	1.4	2295	1	US-08-375-300-3	Sequence 3, Appli
34	47.2	1.4	2295	3	US-09-177-431-3	Sequence 3, Appli
35	47.2	1.4	2295	5	PCT-US95-16930-3	Sequence 3, Appli
36	47.2	1.4	4080	1	US-08-375-300-1	Sequence 1, Appli
37	47.2	1.4	4080	3	US-09-177-431-1	Sequence 1, Appli
38	47.2	1.4	4080	5	PCT-US95-16930-1	Sequence 1, Appli
39	46.8	1.4	1956	3	US-08-559-896B-1	Sequence 1, Appli
40	46.8	1.4	1956	4	US-09-351-794A-1	Sequence 1, Appli
41	46.6	1.4	3399	4	US-09-614-221A-600	Sequence 600, App
42	46	1.4	10304	4	US-09-627-465B-1	Sequence 1, Appli
43	45.8	1.4	546	4	US-09-248-796A-11840	Sequence 11840, A
44	45.8	1.4	832	4	US-09-621-976-2813	Sequence 2813, Ap
45	45.8	1.4	1233	4	US-09-856-930-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-09-023-655-48
; Sequence 48, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,655
; FILING DATE: HERewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1392 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: HMCINOT01
; CLONE: 004700
US-09-023-655-48

Query Match 38.0%; Score 1278.8; DB 4; Length 1392;

Best Local Similarity 94.1%; Pred. No. 0;		Matches 1311; Conservative 0; Mismatches 77; Indels 5; Gaps 2;	
QY	1976	ATTTGAGGAAGTACAATGAGGCCCTACAAATTAATGACACAATTCGAATGATAGTGCCT	2035
Db	1	ATTTGAGGAAGTACAATGAGGCCCTACAAATTAATGACACAATTCGAATGATAGTGCCT	60
QY	2036	ATACTCATCTTGAAACTTTCTATAATGAAGAGAAAGATAAGAAAGTTTGCGAGTCATAGAAG	2095
Db	61	ATACTCATCTTGAAACTTTCTATAATGAAGAGAAAGATAAGAAAGTTTGCGAGTCATANNNN	120
QY	2096	ATGATAGTATGAGGCTGGTGATGATGAGTATTGTGTGATGGTGATGAAGATGAGGATGATT	2155
Db	121	NNNT	180
QY	2156	TAAAGAAACCTTTGAAACTGGATGAAACAGATAGATTCTCATGACTTTTATTTTGAAA	2215
Db	181	TAAAGAAACCTTTGAAACTGGATGAAACAGATAGATTCTCATGACTTTTATTTTGAAA	240
QY	2216	ACAATAAATGTTGAAAGGCTGGCTGAAACCCAGAAATATGAAATGAAAGCTGACCA	2275
Db	241	ACAATAAATGTTGAAAGGCTGGCTGAAACCCAGAAATATGAAATGAAAGCTGACCA	300
QY	2276	AATTAAGAAATACCATAAATGGAGCAATATACTAGGACTGAGGAATCAGCACGAGGAATAA	2335
Db	301	AATTAAGAAATACCATAAATGGAGCAATATACTAGGACTGAGGAATCAGCACGAGGAATAA	360
QY	2336	TCTTTACAAAACACGACAGAGTGCAATATGCGCTTTCCAGTGGATTACTGAAATGAAA	2395
Db	361	TCTTTACAAAACACGACAGAGTGCAATATGCGCTTTCCAGTGGATTACTGAAATGAAA	420
QY	2396	AATTTGCTGAAGTAGAGTCAAAAGCCCAACCATCTGATTGGAGCTGGACACAGCAGTGAGT	2455
Db	421	AATTTGCTGAAGTAGAGTCAAAAGCCCAACCATCTGATTGGAGCTGGACACAGCAGTGAGT	480
QY	2456	TCAAACCCATGACACAGAATGAACAAAAGAAAGTCATTAGTAAATTTGCGACTGGAAAAA	2515
Db	481	TCAAACCCATGACACAGAATGAACAAAAGAAAGTCATTAGTAAATTTGCGACTGGAAAAA	540
QY	2516	TCAATCTGCTTATCGCTACACAGTGGCAGAGAAGGTCTGGATATTAAAGAAATGTAACA	2575
Db	541	TAAATCTGCTTATCGCTACACAGTGGCAGAGAAGGTCTGGATATTAAAGAAATGTAACA	600
QY	2576	TTGTTATCCGTTATGCTCGTCACCAATGAATAGCCATGCTCCAGGCCCGTGGTCGAG	2635
Db	601	TTGTTATCCGTTATGCTCGTCACCAATGAATAGCCATGCTCCAGGCCCGTGGTCGAG	660
QY	2636	CCAGAGCTGATGAGAGCACCTACGTCCTGTTGCTCACAGTGGTTCAGGAGTTATCGAAC	2695
Db	661	CCAGAGCTGATGAGAGCACCTACGTCCTGTTGCTCACAGTGGTTCAGGAGTTATCGAAC	720
QY	2696	ATGAGACAGTTAATGATTTCCGAGAGAAGATGATGTATAAGCTATACATTGTGTTCAA	2755
Db	721	GTGAGACAGTTAATGATTTCCGAGAGAAGATGATGTATAAGCTATACATTGTGTTCAA	780
QY	2756	ATATGAACACGAGGAGTATGCTCATAGATTTTGGAAATACAGATGCAAAAGTATAATGG	2815
Db	781	ATATGAACACGAGGAGTATGCTCATAGATTTTGGAAATACAGATGCAAAAGTATAATGG	840
QY	2816	AAAAGAAATGAAAAACCAAGAGAAATATTGCCAAGCATTAACAAGATAACCCATCACTAA	2875
Db	841	AAAAGAAATGAAAAACCAAGAGAAATATTGCCAAGCATTAACAAGATAACCCATCACTAA	900
QY	2876	TAACTTTCCCTTGCAAAAACCTGCAGTGTGCTAGCCTGTTCTGGGGAAGATATCCATGTAA	2935
Db	901	TAACTTTCCCTTGCAAAAACCTGCAGTGTGCTAGCCTGTTCTGGGGAAGATATCCATGTAA	960
QY	2936	TTGAGAAATGCAATCAGTCATATGACCCAGAAATTAAGAACTTTTACATTGTAAGAG	2995
Db	961	TTGAGAAATGCAATCAGTCATATGACCCAGAAATTAAGAACTTTTACATTGTAAGAG	1020
QY	2996	AAAACAAAGCACTGCAAAAGAAAGTGTGCCGACTATCAATAAATGGTGAATCATCTGCA	3055

Db	1021	AAAAAAAGCACTGCAAAAGAAAGTGTGCCGACTATCAATAAATGGTGAATCATCTGCA	1080
QY	3056	AATGTGCCAGGCTTGGGAAACAATGATGGTGCACAAAGGCTTAGATTGCTTGTCTCA	3115
Db	1081	AATGTGCCAGGCTTGGGAAACAATGATGGTGCACAAAGGCTTAGATTGCTTGTCTCA	1140
QY	3116	AAATAAGGAATTTTGTAGTGGTTTTTCAAAAATAATTCAACAAAGAAACAATACAAAAAGT	3175
Db	1141	AAATAAGGAATTTTGTAGTGGTTTTTCAAAAATAATTCAACAAAGAAACAATACAAAAAGT	1200
QY	3176	GGGTAGAATTACCTATCACATTTCCCAATCTTGACTATTCAGAAATGCTGTTTATTAGTG	3235
Db	1201	GGGTAGAATTACCTATCACATTTCCCAATCTTGACTATTCAGAAATGCTGTTTATTAGTG	1260
QY	3236	ATGAGGATTAGCACCTTGATTGAAGATTCTTTTAAAATACTATCAGTTAAACATTTAATAT	3295
Db	1261	ATGAGGATTAGCACCTTGATTGAAGATTCTTTTAAAATACTATCAGTTAAACATTT--AATA	1318
QY	3296	GATTATGATTAATGTATTCAATTATCTACAGAACTGACATAAGAATC--AATAAAATGA	3352
Db	1319	TGTTATGATTAATGTATTCAITTATGCTACAGAACTGACAATNGAGNCTCAATAAAATGA	1378
QY	3353	TTGTTTTTACTCTG 3365	
Db	1379	TTGTTTTTACTCTG 1391	

RESULT 2
US-08-143-576-7
; Sequence 7, Application US/08143576
; Patent No. 5643761
; GENERAL INFORMATION:
; APPLICANT: Fisher, Paul B.
; APPLICANT: Jiang, Hongping
; TITLE OF INVENTION: METHOD FOR GENERATING A SUBTRACTED
; TITLE OF INVENTION: CDNA LIBRARY AND USES OF THE GENERATED LIBRARY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John P. White, c/o Cooper & Dunham
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/143,576
; FILING DATE: 25-OCT-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 43563/JPW/AKC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 977-9550
; TELEFAX: (212) 664-0525
; TELEX: 422523 COOP UI
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 301 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-143-576-7

Query Match 8.5%; Score 285.4; DB 1; Length 301;

Best Local Similarity 99.3%; Pred. No. 2.4e-67;
Matches 297; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 1592 ACAAAACGAGTATCCCTTCCTCAGATACTGGGACTAACAGCTTACCTGGTGTGGAG 1651
Db 1 ACAAAACGAGTATCCCTTCCTCAGATACTGGGACTAACAGCTTACCTGGTGTGGAG 60
QY 1652 GGGCCACGAAGCCAAAGCTGAAGAACACATTTTAAAACTATGTGCCAATCTTGATG 1711
Db 61 GGGCCACGAAGCCAAAGCTGAAGAACACATTTTAAAACTATGTGCCAATCTTGATG 120
QY 1712 CATTACTATTAAAACTGTTAAAGAAAACCTTGATCAACTGAAAAACCAATACAGGAGC 1771
Db 121 CATTACTATTAAAACTGTTAAAGAAAACCTTGATCAACTGAAAAACCAATACAGGAG- 179
QY 1772 CATGCAAGAAGTTTGCCATTGCGAGATGCAACCCAGAGAAGATCCATTTAAAGAGAACTTC 1831
Db 180 CATGCAAGAAGTTTGCCATTGCGAGATGCAACCCAGAGAAGATCCATTTAAAGAGAACTTC 239
QY 1832 TAGAAATAATGACAAAGGATTCAAACCTTATTGTCAAATGAGTCCAATGTGAGATTTTGA 1890
Db 240 TAGAAATAATGACAAAGGATTCAAACCTTATTGTCAAATGAGTCCAATGTGAGATTTTGA 298

RESULT 3
US-09-221-268D-12
; Sequence 12, Application US/09221268D
; Patent No. 6720408
; GENERAL INFORMATION:
; APPLICANT: Fisher, Paul B.
; APPLICANT: Jiang, Hongping
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES WHICH SPECIFICALLY HYBRIDIZE WITH
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCE ENCODING MDA-7 AND PHARMACEUTICAL COMPOSIT
; FILE REFERENCE: A34534-A-A (070050.1637)
; CURRENT APPLICATION NUMBER: US/09/221,268D
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: 08/316,537
; PRIOR FILING DATE: 1994-09-30
; PRIOR APPLICATION NUMBER: 08/143,576
; PRIOR FILING DATE: 1993-10-27
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 301
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-221-268D-12

Query Match 8.5%; Score 285.4; DB 4; Length 301;
Best Local Similarity 99.3%; Pred. No. 2.4e-67;
Matches 297; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 1592 ACAAAACGAGTATCCCTTCCTCAGATACTGGGACTAACAGCTTACCTGGTGTGGAG 1651
Db 1 ACAAAACGAGTATCCCTTCCTCAGATACTGGGACTAACAGCTTACCTGGTGTGGAG 60
QY 1652 GGGCCACGAAGCCAAAGCTGAAGAACACATTTTAAAACTATGTGCCAATCTTGATG 1711
Db 61 GGGCCACGAAGCCAAAGCTGAAGAACACATTTTAAAACTATGTGCCAATCTTGATG 120
QY 1712 CATTACTATTAAAACTGTTAAAGAAAACCTTGATCAACTGAAAAACCAATACAGGAGC 1771
Db 121 CATTACTATTAAAACTGTTAAAGAAAACCTTGATCAACTGAAAAACCAATACAGGAG- 179
QY 1772 CATGCAAGAAGTTTGCCATTGCGAGATGCAACCCAGAGAAGATCCATTTAAAGAGAACTTC 1831
Db 180 CATGCAAGAAGTTTGCCATTGCGAGATGCAACCCAGAGAAGATCCATTTAAAGAGAACTTC 239
QY 1832 TAGAAATAATGACAAAGGATTCAAACCTTATTGTCAAATGAGTCCAATGTGAGATTTTGA 1890
Db 240 TAGAAATAATGACAAAGGATTCAAACCTTATTGTCAAATGAGTCCAATGTGAGATTTTGA 298

RESULT 4
US-09-799-451-771
; Sequence 771, Application US/09799451
; Patent No. 6783969
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie
; APPLICANT: Xue, Aidong J.
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Ma, Yungqing
; APPLICANT: Yamazaki, Victoria
; APPLICANT: Chen, Rui-hong
; APPLICANT: Wang, Zhiwei
; APPLICANT: Wang, Dunrui
; APPLICANT: Yang, Yonghong
; APPLICANT: Wehrman, Tom
; APPLICANT: Ghosh, Reena
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6783969el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 803
; CURRENT APPLICATION NUMBER: US/09/799,451
; CURRENT FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 948
; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 771
; LENGTH: 2590
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (190)..(2223)
US-09-799-451-771

Query Match 5.3%; Score 179.6; DB 4; Length 2590;
Best Local Similarity 47.6%; Pred. No. 3.9e-38;
Matches 999; Conservative 0; Mismatches 949; Indels 150; Gaps 10;
QY 1048 GAGAAATGTCGACGACGAGCATCCCCGGAGCCAGAACTCCAGCTCAGGCCTTACCAAATG 1107
Db 154 GGGACTGCCCTGGGCAGACCTACTACTAGAGCAGAAATGGAGCTTCGGTCTCTACCAATGG 213
QY 1108 GAAAGTTCGCCAGCCAGCCTTGGAGGGAAGAATATCATCATCTGCCTCCCTACAGGGAGT 1167
Db 214 GAGGTGATCATGCCTGCCCTGGAGGCAAGAATATCATCATCTGGCTGCCACGGGTGCC 273
QY 1168 GGAAAAACCCAGAGTGGCTGTTTACATTGCCAAGGATCACTTAGACAAGAAGAAAAAGCA 1227
Db 274 GGGAAGACCCGGCGGCTGCTTATGTGGCCAAAGCGGCACCTAGAGACTGTGGATGGAGC- 332
QY 1228 TCTGAGCCTGAAAAAGTTATAGTTCTTGTCAATAAGGTACTGCTAGTTGAACAGCTCTTC 1287
Db 333 -----CAAGGTGGTTGTATTGGTCAACAGGGTGACCTGGTGACCCAGC---AT 378
QY 1288 CGCAAGGAGTTCCAACCATTTTGAAGAAATGGTATCGTGTATTGGATTAAGTGGTGAT 1347
Db 379 GGTGAAGAGTTCAGCGGCATGCTGGATGGACGCTGGACCGTGACAAACCTGAGTGGGGAC 438
QY 1348 ACCCAACTGAAAAATATCATTTCCAGAAAGTTGTCAAGTCTCTGTGATATTATTATCAGTACA 1407
Db 439 ATGGGACCACGTGCTGGCTTTGGCCACCTGGCCCGGTGCCATGACCTGCTCATCTGCACA 498
QY 1408 GCTCAAAATCCTTGAAAACTCCCTCTTAAACTTGGAAAAATGGAGAAAGATGCTGGTGTCAA 1467
Db 499 GCAGAGCTTCTGCAGATGGCACTGACCAGCCCCGA-----GGAGGAGGAGCACCTGGAG 552
QY 1468 TTGTCAGACTTTTCCCTCATTTATCATTTGATGAATGTCTATCACACCAACAAAGACGATG 1527

Db 553 CTCACTGTCTTCTCCCTGATCGTGGTGGATGAGTGCCACCACACCGCACCAAGGACACCGCTC 612
Qy 1528 TATAATAACATCATGAGGCATATTTTATGATGCAGAGTTGAAAAACAATAGACTCAAGAAA 1587
Db 613 TACAACGTCTATGAGCCAGTACCT- - - - - AGAAGTAACTCCAGAGG 657
Qy 1588 GAAACAAACAGTGAATCCCTTCCTCAGATACCTGGGACTAACAGCTTCACCTGGTGT 1647
Db 658 GCACA- - - - - GCCGCTACCCAGGTCGTGGTCTCAGAGCTCCCGAGGCACT 705
Qy 1648 GGAGGGGCCACGAAGCAAGCCAAAGCTGAAGAACATTTTAAACATATGTGCCAATCTT 1707
Db 706 GCGGGGCTCCAAACTCGATGGGCCCATCAACCCAGCTCCTGCAGCTCTGTGCCAACTTG 765
Qy 1708 GATGCATTACTATTAAACTGTAAAGAAAACTTGATCAACTGAACCAACCAATACAG 1767
Db 766 GACACGTGTGTCATGTCAACCCAGAACTGCTGCCCCAGCTGCAGGACACAGCCAA 825
Qy 1768 GAGCCATGCAAGAAGTTTGCCATTGCGAGATGCNACCAGAGAAGATCCATTAAAGAGAAA 1827
Db 826 CAGCCTTGCAACAGTACAACCTCTGCCACAGCGCAGCCAGGATCCGTTTGGGCACTTG 885
Qy 1828 CTTCTAGAAATATGACAAGGATTCAAACTTATTGTCAAAATGAGTCCAAATGCA- - - GAT 1884
Db 886 CTGAAGAAGCTCATGGACCAATCCATGACCACCTGGAGATGCTGAGTTGAGCCGGAAA 945
Qy 1885 TTTGGAACTCAACCCCTATGAACAATGSSCCATTCAAATGGAAGAAAAAAGCTGCAAAAAA 1944
Db 946 TTTGGACGCAATGTATGAGCAGCAGGTGTTGAGCTGAGTGGCTGCGGCTTTGGCT 1005
Qy 1945 GGAATCGCAAGAACGTGTTTGTGCAAGCAATTTGAGGAATPACAATGAGGCCCTACAA 2004
Db 1006 GGGCTTCAGGAGCAACGGGTGTATGCGCTTACCTGAGCGGTACAATGACGCGTGTCTC 1065
Qy 2005 ATTAATGACACAATTCGAATGATAGATGCGTATCTCATCTTGAAACTTTCTATAATGAA 2064
Db 1066 ATCCATGACACCGTCCGCGCGTGGATGCGCTGGCTGCGCTGAGGATTTCTATCACAGG 1125
Qy 2065 GAGAAAGATAAGAGTTTGCAGTCAATAGAGATGATGATGAGGGTGGTATGATGAG 2124
Db 1126 GAGACGTCACTAAACCCAGATCCTGTGTG- - - - - 1156
Qy 2125 TATTGTGATGGTATGAAGATGAGGATGATTTAAAGAAACCTTTGAAACTGGATGAACA 2184
Db 1157 - - - - - CC 1158
Qy 2185 GATAGATTTCTCATGACTTTATTTTGAACAATAAATTTGAAAAGGCTGGCTGAA 2244
Db 1159 GAGCGCGGCTGCTGGCCCTGTTTCTGATGACCGGAAGATGAGCTGGCCCACTTGGC- - - A 1215
Qy 2245 AACCCAGAAATATGAAAATGAAAAGCTGACCAATTAAGAAATACCAATATGGAGCAATAT 2304
Db 1216 ACTCATGGCCAGAGAAATCCAAAACCTGGAGATGCTGGAAGATCTCTGCAAGGCAATTC 1275
Qy 2305 ACTAGGACTGAGGAATCAGCAGCAGGAATAATCTTTACAAAAACACACAGAGTGCAAT 2364
Db 1276 AGTAGCTCT- - - AACAGCCCTCGGGGTATCATCTTCAACCCGACCCGCAAAAGCGCACAC 1332
Qy 2365 GCGCTTTCCAGTGGATTACTGAAAATGAAAAATTTGCTGAAGTAGGAGTCAAGCCAC 2424
Db 1333 TCCCTCTGCTGCTGGCTCCAGCAGCAGCAGGCGCTGCAGACTGTGGACATCCGGGCCAG 1392
Qy 2425 CATCTGATTGGAGCTGGACACAGCAGTGAGTTCAACCCCATGACACAGAAATGAACAAAA 2484
Db 1393 CTACTGATTGGGCTGGAAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1452
Qy 2485 GAAGTCATTACTAAATTTCCGCACTGGAAAAATCAATCTGCTTATCGTACACAGTGGCA 2544
Db 1453 GAAGTGATCAGAAAGTTCCAAGATGGAACCTTGAACCTTCTGGTGGCCACAGTGTGGC 1512
Qy 2545 GAAGAAGGCTGGATATTAAGAATGTAACATTGTTATCCGTTATGGTCTCTCTCACCAAT 2604
Db 1513 GAGGAGGGGCTGGACATCCACATGCAATGTGGTGGTGTATGGCTCTTGACCAAT 1572

Qy 2605 GAAATAGCCATGGTCCAGGCCCGTGGTCGAGCCAGAGCTGATGAGAGCACCTACGTCCTG 2664
Db 1573 GAAATCTCATGGTCCAGGCCAGGGCCGTCGCCGGCCGATCAGAGTGTATACGCGTTT 1632
Qy 2665 GTTGTCTACAGTGGTTCAGGAGTTATCGAACATGAGACAGTTAATGATTTCCGAGAGAAG 2724
Db 1633 GTAGCAACTGAAGTAGCCGGAGCTGAAGCGGAGCTGATCAACGAGGCGCTGGAGACG 1692
Qy 2725 ATGATGTATAAAGCTATACATTTGTTTCAAAATATGAAAACAGAGGAGTATGCTCATAAG 2784
Db 1693 CTGATGGAGCAGCAGTGGTGTCTGTGCAGAAAATGGACCGAGCCGAGTACCAGGCCAAG 1752
Qy 2785 ATTTTGAATTACAGATGCAAGATTAATGGAAGAAAAATGAAAACCAAGAGAAATATT 2844
Db 1753 ATCCGGGATCTGCAGCAGGACGCTTGACCAAGCGGGCGCCAGCAGCCAGCGGAG 1812
Qy 2845 GCCAAGCATTAAGAATAAACCCTACTAATAACTTTTCTTTGCAAAAACTGCAGTGTG 2904
Db 1813 AACCAAGCGCAGCAGTTCCAGTGGAGCAGCTGCAGCTACTCTGCATCAACTGCATGGTG 1872
Qy 2905 CTAGCCTGTTCTGGGGAAGATATCCATGTAATTGAGAAAATGCATCACGTCAATATGACC 2964
Db 1873 GCTGTGGCCATGGCAGCAGACCTGCGGAAGTGGAGGACCCACCATGTCAATGTGAAC 1932
Qy 2965 CCAGAATTCAAGGAACTTTACATTTGTAAGAGAAAAACAAGCACTGCAAAAGAGTGTGCC 3024
Db 1933 CCCAACTTCTCGAACTACTATAATGTCTCCAGGATCCTGTGGTCAACAACAAGTCTTC 1992
Qy 3025 GACTATCAAAATAAATGGTGAATCATCTGC- - - - - AAATGTGGCCAGGCTTGGGAACA 3078
Db 1993 AAGACTGGAAGCCTGGGGTGTATCATCAGCTGCAGGAATGTGGGGAGGTCTGGGTCTG 2052
Qy 3079 ATGATGGTGCAAAAGGCTTAGATTTGCTTGTCTCAAAATAAGGAATTTTGTAGTGG 3136
Db 2053 CAGATGATCTACAAGTCAAGTGAAGTGCAGTGCAGTCTCAAAAGTCCGACGATGCTGCTGG 2110

RESULT 5

US-08-143-576-6
; Sequence 6, Application US/08143576
; Patent No. 5643761
; GENERAL INFORMATION:
; APPLICANT: Fisher, Paul B.
; APPLICANT: Jiang, Hongping
; TITLE OF INVENTION: METHOD FOR GENERATING A SUBTRACTED
; TITLE OF INVENTION: CDNA LIBRARY AND USES OF THE GENERATED LIBRARY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John P. White, c/o Cooper & Dunham
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/143,576
; FILING DATE: 25-OCT-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 43563/JPW/AKC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 977-9550
; TELEFAX: (212) 664-0525
; TELEX: 422523 COOP UI
; INFORMATION FOR SEQ ID NO: 6:


```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 193 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-143-576-6

Query Match          3.7%; Score 124.6; DB 1; Length 193;
Best Local Similarity 96.5%; Pred. No. 6.2e-24;
Matches 138; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 3007 CTGCAAAAGAGTGTGCCGACTATCAATAAATGGTGAAATCATCTGCAAAATGTGGCCAG 3066
Db 1 CTGCAAAAGAGTGTGCCGACTAT-AAATAAATGGTGAAATCATCTGCAAAATGTGGCCAG 59

QY 3067 GCTTGGGGAACAATGATGGTGCACAAAGGCTTAGATTTGCTCTCAAAATAAGGAAT 3126
Db 60 GCTTGGGGAACAATGATGGTGCACAAAGGCTTAGATTTGCTCTCAAAATAAGGAAT 119

QY 3127 TTTGTAGTGGTTTCAAAAATAA 3149
Db 120 TTTGTAGTGGTTTCAAAAATAATCA 142

RESULT 6
US-09-221-268D-11
; Sequence 11, Application US/09221268D
; Patent No. 6720408
; GENERAL INFORMATION:
; APPLICANT: Fisher, Paul B.
; APPLICANT: Jiang, Hongping
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES WHICH SPECIFICALLY HYBRIDIZE WITH
; FILE REFERENCE: A34534-A-A (070050.1637)
; CURRENT APPLICATION NUMBER: US/09/221,268D
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: 08/316,537
; PRIOR FILING DATE: 1994-09-30
; PRIOR APPLICATION NUMBER: 08/143,576
; PRIOR FILING DATE: 1993-10-27
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 193
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-221-268D-11

Query Match          3.7%; Score 124.6; DB 4; Length 193;
Best Local Similarity 96.5%; Pred. No. 6.2e-24;
Matches 138; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 3007 CTGCAAAAGAGTGTGCCGACTATCAATAAATGGTGAAATCATCTGCAAAATGTGGCCAG 3066
Db 1 CTGCAAAAGAGTGTGCCGACTAT-AAATAAATGGTGAAATCATCTGCAAAATGTGGCCAG 59

QY 3067 GCTTGGGGAACAATGATGGTGCACAAAGGCTTAGATTTGCTCTCAAAATAAGGAAT 3126
Db 60 GCTTGGGGAACAATGATGGTGCACAAAGGCTTAGATTTGCTCTCAAAATAAGGAAT 119

QY 3127 TTTGTAGTGGTTTCAAAAATAA 3149
Db 120 TTTGTAGTGGTTTCAAAAATAATCA 142

RESULT 7
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
```

```

; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-F1s
US-08-232-463-14

Query Match          2.1%; Score 69.4; DB 1; Length 7218;
Best Local Similarity 4.8%; Pred. No. 5e-08;
Matches 19; Conservative 230; Mismatches 146; Indels 0; Gaps 0;

QY 1876 ATGTCAGATTTTGGAACTCAACCTATGAACAATGGGCCATTCAAATGGAAAAAAGCT 1935
Db 1451 ATAGAAGAAATTGGTACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1392

QY 1936 GCAAAAAAGGAAATCGAAAGAACGTGTTTGTGCAGAACATTTGAGGAAGTACAATGAG 1995
Db 1391 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1332

QY 1996 GCCCTACAAATTAATGACACAATTCGAATGATAGATCGGTACTCATCTTGAAACTTC 2055
Db 1331 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1272

QY 2056 TATAATGAAGAGAAAGATAAGAGTTTGCAGTCATAGAAGATGATGATGAGGTGCT 2115
Db 1271 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1212

QY 2116 GATGATGAGTATTGTGATGGTGATGAAGATGAGGATGATTAAAGAAACCTTTGAAACTG 2175
Db 1211 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1152

QY 2176 GATGAAACAGATAGATTTCTCATGACTTTATTTTTTGAACAAATAAAATGTTGAAAGG 2235
Db 1151 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1092

QY 2236 CTGGCTGAAACCAGAAATATGAAAAATGAAAAGCT 2270
```


244 AATGCTT--GGTGTAAATGAAAAAGAACAAATAAACATGTACAAAGAAAAAATGATGAAT 187
2734 AAAGCTATACATTGTGTTCAAAAATATGAACCAGAGGAGTATGCTCATAAGATTTTGGAA 2793
186 GACCTTTATTTACGCTTCAGACATGGGACGAGCAGTATTTAGGGGAAAAAGATTCTGCAT 127
2794 TTACAGATGCAAAAGTATAATGGAAGAAAGAAATCAAAAACCAAGAGAAATATTGCCAAGCAT 2853
126 ATACAGACTCATGAAAAATTCATCAGAGATAGTCAAGAAAAACCAAAACCTGTACCTGAT 67
2854 TACAAGAAATAA 2864
66 AAGGAAAAATAA 56

Search completed: January 27, 2005, 17:57:24
Job time : 298 secs

2555 TGGATATTAAAGAATGTAACATTGTTATCCGTTATG-GTCTCGTCACCAATGAAATAGCC 2613
364 TTGACATTGCACAGTGCAATCTTGTTCATCCTTCATGAGTAATGTGGGCAATGTCATCAA 305
2614 ATGGTCCAGGCCGTTGTCGAGCCAGAGCTGATGAGAGCACCTACGTCCTCGTTGCTCAC 2673
304 ATGATCCAAACAGAGGCAGAGGGAAGAGCAAGAGGTAGCAAGTGCTTCTTCTGACTAGT 245
2674 AGTGGTTCAGGAGTTATCGAACATGAGACAGTTAATGATTTCCGAGAGAAGATGATGTAT 2733
244 AATGCTT--GGTGTAAATGAAAAAGAACAAATAAACATGTACAAAGAAAAAATGATGAAT 187
2734 AAAGCTATACATTGTGTTCAAAAATATGAACCAGAGGAGTATGCTCATAAGATTTGGAA 2793
186 GACCTTTATTTACGCTTCAGACATGGGACGAGCAGTATTTAGGGAAAAAGATTCTGCAT 127
2794 TTACAGATGCAAAAGTATAATGGAAGAAAGAAATGAAAAACCAAGAGAAATATTGCCAAGCAT 2853
126 ATACAGACTCATGAAAAATTCATCAGAGATAGTCAAGAAAAACCAAAACCTGTACCTGAT 67
2854 TACAAGAAATAA 2864
66 AAGGAAAAATAA 56

RESULT 15
US-09-433-826B-46/c
; Sequence 46, Application US/09433826B
; Patent No. 6579973
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqi
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.470C4
; CURRENT APPLICATION NUMBER: US/09/433,826B
; CURRENT FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 474
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 46
; LENGTH: 371
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-433-826B-46

Query Match 1.6%; Score 54.6; DB 4; Length 371;
Best Local Similarity 52.7%; Pred. No. 8.2e-05;
Matches 164; Conservative 0; Mismatches 144; Indels 3; Gaps 2;
QY 2555 TGGATATTAAAGAATGTAACATTGTTATCCGTTATG-GTCTCGTCACCAATGAAATAGCC 2613
Db 364 TTGACATTGCACAGTGCAATCTTGTTCATCCTTCATGAGTAATGTGGGCAATGTCATCAA 305
QY 2614 ATGGTCCAGGCCGTTGTCGAGCCAGAGCTGATGAGAGCACCTACGTCCTGGTTGCTCAC 2673
Db 304 ATGATCCAAACAGAGGCAGAGGGAAGAGCAAGAGGTAGCAAGTGCTTCTTCTGACTAGT 245
QY 2674 AGTGGTTCAGGAGTTATCGAACATGAGACAGTTAATGATTTCCGAGAGAAGATGATGTAT 2733

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 27, 2005, 15:18:52 ; Search time 1698 Seconds
(without alignments)
11386.875 Million cell updates/sec

Title: US-09-515-363C-1
Perfect score: 3365
Sequence: 1 gcgcgcgcgcctgagagccc.....aaaatgattgtttactctg 3365

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4300275 seqs, 2872944193 residues

Total number of hits satisfying chosen parameters: 8600550

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
20: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
21: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	3365	100.0	3365	14	US-10-055-475-1	Sequence 1, Appli
2	3365	100.0	3365	14	US-10-228-897-1	Sequence 1, Appli
3	3365	100.0	3380	17	US-10-755-889-405	Sequence 405, App
4	3361.8	99.9	3365	14	US-10-055-475-8	Sequence 8, Appli
5	3360.2	99.9	3928	14	US-10-198-846-13042	Sequence 13042, A
6	3358.6	99.8	3627	14	US-10-055-475-4	Sequence 4, Appli
7	3339.2	99.2	3668	18	US-10-723-860-5057	Sequence 5057, Ap
8	3338.8	99.2	3372	16	US-10-275-822A-1	Sequence 1, Appli
9	1766	52.5	1776	15	US-10-094-749-828	Sequence 828, App
10	1435.6	42.7	1443	16	US-10-275-822A-10	Sequence 10, Appl
11	1278.8	38.0	1392	16	US-10-641-643-48	Sequence 48, Appl
12	1254.6	37.3	1284	16	US-10-275-822A-7	Sequence 7, Appli

13	577.2	17.2	609	15	US-10-106-698-811	Sequence 811, App
14	577.2	17.2	609	16	US-10-264-049-69	Sequence 69, Appl
15	556	16.5	6406	14	US-10-228-897-25	Sequence 25, Appl
16	554	16.5	1036	14	US-10-055-475-3	Sequence 3, Appli
17	554	16.5	1036	14	US-10-228-897-3	Sequence 3, Appli
c 18	507.2	15.1	595	9	US-09-864-761-12563	Sequence 12563, A
c 19	455.2	13.5	499	9	US-09-864-761-29164	Sequence 29164, A
c 20	453	13.5	453	15	US-10-029-386-24101	Sequence 24101, A
21	419.4	12.5	457	10	US-09-918-995-12953	Sequence 12953, A
22	376.4	11.2	458	10	US-09-907-907A-32	Sequence 32, Appl
c 23	342	10.2	501	15	US-10-029-386-10395	Sequence 10395, A
24	330.4	9.8	403	14	US-10-198-846-5938	Sequence 5938, Ap
25	285.4	8.5	301	15	US-10-417-827-12	Sequence 12, Appl
26	281.4	8.4	292	9	US-09-864-761-17369	Sequence 17369, A
27	252.8	7.5	460	9	US-09-864-761-577	Sequence 577, App
c 28	238.8	7.1	465	18	US-10-723-860-354	Sequence 354, App
c 29	229.6	6.8	554	9	US-09-864-761-15269	Sequence 15269, A
c 30	217	6.4	217	9	US-09-864-761-31791	Sequence 31791, A
31	194.4	5.8	450	9	US-09-864-761-15558	Sequence 15558, A
32	179.6	5.3	2590	16	US-10-302-172-771	Sequence 771, App
c 33	174	5.2	386	9	US-09-864-761-14038	Sequence 14038, A
34	167	5.0	167	9	US-09-864-761-32068	Sequence 32068, A
35	153	4.5	164	16	US-10-275-822A-3	Sequence 3, Appli
36	150.2	4.5	2250	16	US-10-108-260A-2072	Sequence 2072, Ap
37	148.6	4.4	3065	15	US-10-172-118-1472	Sequence 1472, Ap
38	148.6	4.4	3065	16	US-10-342-887-1472	Sequence 1472, Ap
39	148.6	4.4	3065	17	US-10-755-889-233	Sequence 233, App
40	148.6	4.4	3065	18	US-10-370-715B-7	Sequence 7, Appli
41	147	4.4	3258	13	US-10-044-090-631	Sequence 631, App
c 42	145	4.3	145	9	US-09-864-761-30598	Sequence 30598, A
43	124.6	3.7	193	15	US-10-417-827-11	Sequence 11, Appl
44	122.4	3.6	486	10	US-09-918-995-32015	Sequence 32015, A
45	109.6	3.3	3184	16	US-10-264-049-128	Sequence 128, App

ALIGNMENTS

RESULT 1
US-10-055-475-1
; Sequence 1, Application US/10055475
; Publication No. US20030022855A1
; GENERAL INFORMATION:
; APPLICANT: Fisher, Paul B.
; APPLICANT: Kang, Dong-Chul
; APPLICANT: Gopalkrishnan, Rahul V.
; TITLE OF INVENTION: USE OF MDA-5 AS AN ANTIVIRAL AND
; TITLE OF INVENTION: ANTIPROLIFERATIVE AGENT
; FILE REFERENCE: A34614-A-PCT-USA-A (070050.1921)
; CURRENT APPLICATION NUMBER: US/10/055,475
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: PCT/US01/06960
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 09/515,363
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3365
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-055-475-1

Query Match	100.0%	Score 3365;	DB 14;	Length 3365;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 3365;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	GC GCGCGCGCCTGAGAGCCCTGTGGACAACCTCGTCATTGTCAGGCACAGACGGGTAGAC	60	
Db	1	GC GCGCGCGCCTGAGAGCCCTGTGGACAACCTCGTCATTGTCAGGCACAGACGGGTAGAC	60	
Qy	61	CCTGCTTCTTAAGTGGCAGCGGCACGCGGCACGCACATTCACCTGTCCCGCAGACAA	120	

Db	61	CCTGCTTCTCTAAGTGGGACGGGACAGCGGCACGCACATTTCACTGTGTCCTCCGCAGACAA	120
Qy	121	CAGCACCATCTGCTTGGGAAACCTCTCCCTTCTCTGAGAAAGAAAGATGTCGAATGGG	180
Db	121	CAGCACCATCTGCTTGGGAAACCTCTCCCTTCTCTGAGAAAGAAAGATGTCGAATGGG	180
Qy	181	TATTCCACAGACGAGAAATTTCCGCTATCTCATCTCGTGTCTTCAGGGCCAGGGTGAAATG	240
Db	181	TATTCCACAGACGAGAAATTTCCGCTATCTCATCTCGTGTCTTCAGGGCCAGGGTGAAATG	240
Qy	241	TACATCCAGGTGGAGCCTGTGCTGGACTACCTTGACCTTTCTGCCTGCAAGGTGAAGGAG	300
Db	241	TACATCCAGGTGGAGCCTGTGCTGGACTACCTTGACCTTTCTGCCTGCAAGGTGAAGGAG	300
Qy	301	CAGATTCAGAGGACAGTCGCCACCTCCGGGAACATGCAGGCAGTTGAACTGCTGTGAGC	360
Db	301	CAGATTCAGAGGACAGTCGCCACCTCCGGGAACATGCAGGCAGTTGAACTGCTGTGAGC	360
Qy	361	ACCTTGGAGAGGGAGTCTGGCACCTTGGTGTGGACTCGGGAATTCGTGGAGGCCCTCCGG	420
Db	361	ACCTTGGAGAGGGAGTCTGGCACCTTGGTGTGGACTCGGGAATTCGTGGAGGCCCTCCGG	420
Qy	421	AGAACCGGCAGCCCTCTGGCCCGCCGTACATGAACCTTGAGCTCACGGACTTGCCCTCT	480
Db	421	AGAACCGGCAGCCCTCTGGCCCGCCGTACATGAACCTTGAGCTCACGGACTTGCCCTCT	480
Qy	481	CCATCGTTTGAGAACGCTCATGATGAATATCTCCAACTGCTGAACCTCCTTCAGCCCACT	540
Db	481	CCATCGTTTGAGAACGCTCATGATGAATATCTCCAACTGCTGAACCTCCTTCAGCCCACT	540
Qy	541	CTGGTGGACAAGCTTCTAGTTAGAGACGCTCTGGATAAGTGCATGGAGGAGGAAGTGTG	600
Db	541	CTGGTGGACAAGCTTCTAGTTAGAGACGCTCTGGATAAGTGCATGGAGGAGGAAGTGTG	600
Qy	601	ACAATTTGAAGACAGAAACCGGATTGCTGTCAGAAACCAATGGAAATGAATCAGGTGTA	660
Db	601	ACAATTTGAAGACAGAAACCGGATTGCTGTCAGAAACCAATGGAAATGAATCAGGTGTA	660
Qy	661	AGAGAGCTACTAAAAAGGATTGTGCAGAAAGAAACCTGGTTCTCTGCAATTTCTGAAATGT	720
Db	661	AGAGAGCTACTAAAAAGGATTGTGCAGAAAGAAACCTGGTTCTCTGCAATTTCTGAAATGT	720
Qy	721	CTTCGTCAAAACAGGAAACCAATGAACTTGTCCAAGAGTTAACAGGCTCTGATTGCTCAGAA	780
Db	721	CTTCGTCAAAACAGGAAACCAATGAACTTGTCCAAGAGTTAACAGGCTCTGATTGCTCAGAA	780
Qy	781	AGCAATGCAGAGATTGAGAAATTTATCACAAGTTGATGGTCTCAAGTGAAGAGCAACTT	840
Db	781	AGCAATGCAGAGATTGAGAAATTTATCACAAGTTGATGGTCTCAAGTGAAGAGCAACTT	840
Qy	841	CTTTCAACCACAGTTTCAGCCAAATCTGGAGAAAGAGGTCTGGGGCATGGAGAAATACTCA	900
Db	841	CTTTCAACCACAGTTTCAGCCAAATCTGGAGAAAGAGGTCTGGGGCATGGAGAAATACTCA	900
Qy	901	TCAGAATCATCTTTTGCAGATTCTTCTGTAGTTTCAGAAATCAGACACAAAGTTTGGCAGAA	960
Db	901	TCAGAATCATCTTTTGCAGATTCTTCTGTAGTTTCAGAAATCAGACACAAAGTTTGGCAGAA	960
Qy	961	GGAAGTGTAGCTGCTTAGATGAAAGTCTTGGACATAACAGCAACATGGGCAGTGAATTC	1020
Db	961	GGAAGTGTAGCTGCTTAGATGAAAGTCTTGGACATAACAGCAACATGGGCAGTGAATTC	1020
Qy	1021	GGCACCATGGGAAGTGATTCAGATGAAGAGAAATGTGGCAGCAAGAGCATCCCCGGAGCCA	1080
Db	1021	GGCACCATGGGAAGTGATTCAGATGAAGAGAAATGTGGCAGCAAGAGCATCCCCGGAGCCA	1080
Qy	1081	GAACTCCAGCTCAGGCCCTTACCAATGGAAGTTGCCACGACGCTTGGAGGGAAGAAAT	1140
Db	1081	GAACTCCAGCTCAGGCCCTTACCAATGGAAGTTGCCACGACGCTTGGAGGGAAGAAAT	1140
Qy	1141	ATCATCATCTGCCTCCCTACAGGGAGTGGAAAAACCAAGATGGCTGTTTACATTTGCCAAG	1200
Db	1141	ATCATCATCTGCCTCCCTACAGGGAGTGGAAAAACCAAGATGGCTGTTTACATTTGCCAAG	1200

Qy	1201	GATCACTTAGACAAGAAAAAGCACTGTAGCCTGGAAAAAGTTATAGTTCTTTGTCAAT	1260
Db	1201	GATCACTTAGACAAGAAAAAGCACTGTAGCCTGGAAAAAGTTATAGTTCTTTGTCAAT	1260
Qy	1261	AAGTACTGTAGTTGAACAGCTCTTCCGCAAGGAGTTCCAACCAATTTTGAAGAAATGG	1320
Db	1261	AAGTACTGTAGTTGAACAGCTCTTCCGCAAGGAGTTCCAACCAATTTTGAAGAAATGG	1320
Qy	1321	TATCGTGTATTGGATTAAAGTGGTGATACCCCAACTGAAAAATATCATTTCCAGAAGTTGTC	1380
Db	1321	TATCGTGTATTGGATTAAAGTGGTGATACCCCAACTGAAAAATATCATTTCCAGAAGTTGTC	1380
Qy	1381	AAGTCTGTGATATTATTTATCAGTACAGCTCAAAATCCTTGAAAACTCCCTCTTTAAACTTG	1440
Db	1381	AAGTCTGTGATATTATTTATCAGTACAGCTCAAAATCCTTGAAAACTCCCTCTTTAAACTTG	1440
Qy	1441	GAAAAATGGAGAAAGATGCTGGTGTTCAAATTGTTCAGACTTTTCCCTCATTTATCATTTGATGAA	1500
Db	1441	GAAAAATGGAGAAAGATGCTGGTGTTCAAATTGTTCAGACTTTTCCCTCATTTATCATTTGATGAA	1500
Qy	1501	TGTCATCACACCAACAAAGAAAGCAGTGTATAATAACATCATGAGGCATTTATTTGATGCAG	1560
Db	1501	TGTCATCACACCAACAAAGAAAGCAGTGTATAATAACATCATGAGGCATTTATTTGATGCAG	1560
Qy	1561	AAGTTGAAAAACAATAGACTCAAGAAAGAAAAACAACCAAGTATTCCTCTCAGATA	1620
Db	1561	AAGTTGAAAAACAATAGACTCAAGAAAGAAAAACAACCAAGTATTCCTCTCAGATA	1620
Qy	1621	CTGGGACTAACAGCTTCACCTGGTGTTCAGGGGCCACGAAGCAAGCCAAAGCTGAAGAA	1680
Db	1621	CTGGGACTAACAGCTTCACCTGGTGTTCAGGGGCCACGAAGCAAGCCAAAGCTGAAGAA	1680
Qy	1681	CACATTTTAAAACTATGTGCCAATCTTGATGCAATTTACTATTAAAACTGTTTAAAGAAAAAC	1740
Db	1681	CACATTTTAAAACTATGTGCCAATCTTGATGCAATTTACTATTAAAACTGTTTAAAGAAAAAC	1740
Qy	1741	CTTGATCAACTGAAAAACCAAAATACAGAGGCCCATGCAAGAGTTTGCCATTGCGAGATGCA	1800
Db	1741	CTTGATCAACTGAAAAACCAAAATACAGAGGCCCATGCAAGAGTTTGCCATTGCGAGATGCA	1800
Qy	1801	ACCAGAGAAGATCCATTTAAAGAGAAACTTCTAGAAATATGCAAGGATTCAAACTTAT	1860
Db	1801	ACCAGAGAAGATCCATTTAAAGAGAAACTTCTAGAAATATGCAAGGATTCAAACTTAT	1860
Qy	1861	TGTCAAATGAGTCCAAATGTGAGTTTGGAACTCAACCCCTATGAACAATGGGCCATTCAA	1920
Db	1861	TGTCAAATGAGTCCAAATGTGAGTTTGGAACTCAACCCCTATGAACAATGGGCCATTCAA	1920
Qy	1921	ATGGAATAAAAGCTGCAAAAAAAGGAAATCGCAAGAACGTTGTTGTCAGAACATTTG	1980
Db	1921	ATGGAATAAAAGCTGCAAAAAAAGGAAATCGCAAGAACGTTGTTGTCAGAACATTTG	1980
Qy	1981	AGGAAGTACAATGAGGCCCTACAAATTAATGACACAATTCGAATGATAGATGCGTATACT	2040
Db	1981	AGGAAGTACAATGAGGCCCTACAAATTAATGACACAATTCGAATGATAGATGCGTATACT	2040
Qy	2041	CATCTTGAAACTTTCTATAATGAAGAGAAAGATAAGAAAGTTTGCAGTCATAGAGATGAT	2100
Db	2041	CATCTTGAAACTTTCTATAATGAAGAGAAAGATAAGAAAGTTTGCAGTCATAGAGATGAT	2100
Qy	2101	AGTGATGAGGGTGGTGATGATGAGTATTGTGATGGTGATGAAGATGAGGATGATTTAAAG	2160
Db	2101	AGTGATGAGGGTGGTGATGATGAGTATTGTGATGGTGATGAAGATGAGGATGATTTAAAG	2160
Qy	2161	AAACCTTTGAAACTGGATGAACAGATAGATTTCTCATGACTTTATTTTGAACAAT	2220
Db	2161	AAACCTTTGAAACTGGATGAACAGATAGATTTCTCATGACTTTATTTTGAACAAT	2220
Qy	2221	AAAATGTTGAAAAAGGCTGGCTGAAAAACCCAGAAATGAAAAATGAAAGCTGACCAATTA	2280
Db	2221	AAAATGTTGAAAAAGGCTGGCTGAAAAACCCAGAAATGAAAAATGAAAGCTGACCAATTA	2280

QY 2281 AGAAATACCATATGAGCAATATAGGACTGAGGAATCAGCAGGAGGAATAATCTTT 2340
Db |||||
QY 2281 AGAAATACCATATGAGCAATATAGGACTGAGGAATCAGCAGGAGGAATAATCTTT 2340
Db |||||
QY 2341 ACAAAAACGACAGAGTGCATATGCGCTTTCCAGTGGATTACTGAAAAATGAAAAATTT 2400
Db |||||
QY 2401 GCTGAAGTAGGAGTCAAAGCCCACCATCTGATTGGAGCTGGACAGCAGTGAAGTTCAAA 2460
Db |||||
QY 2401 GCTGAAGTAGGAGTCAAAGCCCACCATCTGATTGGAGCTGGACAGCAGTGAAGTTCAAA 2460
Db |||||
QY 2461 CCCATGACACAGAAAGAGAGAGTCAATAGTAAATTCGCACTGGAATAATCAAT 2520
Db |||||
QY 2461 CCCATGACACAGAAAGAGAGAGTCAATAGTAAATTCGCACTGGAATAATCAAT 2520
Db |||||
QY 2521 CTGCTTATCGCTACACAGTGGCAGAGAGGCTCTGGATATTAAAGAATGTAACATTGTT 2580
Db |||||
QY 2521 CTGCTTATCGCTACACAGTGGCAGAGAGGCTCTGGATATTAAAGAATGTAACATTGTT 2580
Db |||||
QY 2581 ATCCGTTATGGTCTCGTCACCAATGAAATAGCCATGGTCCAGGCCGTGGTCGAGCCAGA 2640
Db |||||
QY 2581 ATCCGTTATGGTCTCGTCACCAATGAAATAGCCATGGTCCAGGCCGTGGTCGAGCCAGA 2640
Db |||||
QY 2641 GCTGATGAGAGCCTACGTCCTGGTTGCTCACAGTGGTTCAGGAGTTATCGAACATGAG 2700
Db |||||
QY 2641 GCTGATGAGAGCCTACGTCCTGGTTGCTCACAGTGGTTCAGGAGTTATCGAACATGAG 2700
Db |||||
QY 2701 ACAGTTAATGATTTCCGAGAGAGATGATGTATAAAGCTATACATTTGTTCAAAAATATG 2760
Db |||||
QY 2701 ACAGTTAATGATTTCCGAGAGAGATGATGTATAAAGCTATACATTTGTTCAAAAATATG 2760
Db |||||
QY 2761 AAACCAGAGGAGTATGCTCATAGATTTTGGAAATTACAGATGCAAGTATAATGGAAAG 2820
Db |||||
QY 2761 AAACCAGAGGAGTATGCTCATAGATTTTGGAAATTACAGATGCAAGTATAATGGAAAG 2820
Db |||||
QY 2821 AAAATGAAAACCAAGAGAAATATTGCCAAGCATTACAAGATAACCCATCACTAATAACT 2880
Db |||||
QY 2821 AAAATGAAAACCAAGAGAAATATTGCCAAGCATTACAAGATAACCCATCACTAATAACT 2880
Db |||||
QY 2881 TTCCCTTTGCAAAAACGTCAGTGTGCTAGCTGTTCTGGGGAAGATATCCATGTAATTGAG 2940
Db |||||
QY 2881 TTCCCTTTGCAAAAACGTCAGTGTGCTAGCTGTTCTGGGGAAGATATCCATGTAATTGAG 2940
Db |||||
QY 2941 AAAATGCATCACGTCAAATATGACCCCAAGAAATCAAGGAACTTTACATTTGAAGAGAAAC 3000
Db |||||
QY 2941 AAAATGCATCACGTCAAATATGACCCCAAGAAATCAAGGAACTTTACATTTGAAGAGAAAC 3000
Db |||||
QY 3001 AAAGCACTGCAAAAGAAAGTGTGCCGACTATCAAAATAAATGGTGAATCATCTGCAATGT 3060
Db |||||
QY 3001 AAAGCACTGCAAAAGAAAGTGTGCCGACTATCAAAATAAATGGTGAATCATCTGCAATGT 3060
Db |||||
QY 3061 GGCCAGGCTTGGGGAACAATGATGGTGCAAAAGGCTTAGATTGCTTGTCTCAAATA 3120
Db |||||
QY 3061 GGCCAGGCTTGGGGAACAATGATGGTGCAAAAGGCTTAGATTGCTTGTCTCAAATA 3120
Db |||||
QY 3121 AGGAATTTTGTAGTGGTTTCAAATAAATTAATCAAAAGAAACAATACAAAAAGTGGGTA 3180
Db |||||
QY 3121 AGGAATTTTGTAGTGGTTTCAAATAAATTAATCAAAAGAAACAATACAAAAAGTGGGTA 3180
Db |||||
QY 3181 GAATTACCTATCACATTTCCCAATCTTGACTATTTCAGAAATGCTGTTTATTTAGTGATGAG 3240
Db |||||
QY 3181 GAATTACCTATCACATTTCCCAATCTTGACTATTTCAGAAATGCTGTTTATTTAGTGATGAG 3240
Db |||||
QY 3241 GATTAGCACTTGATTGAAGATTCTTTTAAATACTATCAGTTAAACATTTAATATGATTA 3300
Db |||||
QY 3241 GATTAGCACTTGATTGAAGATTCTTTTAAATACTATCAGTTAAACATTTAATATGATTA 3300
Db |||||
QY 3301 TGATTAATGTATTCAATTATGCTACAGAACTGACATAAGAATCAATAAAATGATTGTTT 3360
Db |||||
QY 3301 TGATTAATGTATTCAATTATGCTACAGAACTGACATAAGAATCAATAAAATGATTGTTT 3360
Db |||||
QY 3361 CTCTG 3365

Db 3361 CTCTG 3365
|||
RESULT 2
US-10-228-897-1
; Sequence 1, Application US/10228897
; Publication No. US20030092043A1
; GENERAL INFORMATION:
; APPLICANT: Fisher, Paul B.
; APPLICANT: Kang, Dong-Chul
; APPLICANT: Gopalrishnan, Rahul V.
; TITLE OF INVENTION: MELANOMA DIFFERENTIATION ASSOCIATED
; TITLE OF INVENTION: GENE-5 AND PROMOTER AND USES THEREOF
; FILE REFERENCE: A34614-A-PCT-USA (070050.2121)
; CURRENT APPLICATION NUMBER: US/10/228,897
; CURRENT FILING DATE: 2002-08-26
; PRIOR APPLICATION NUMBER: PCT/US01/06960
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 09/515,363
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3365
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-228-897-1
Query Match 100.0%; Score 3365; DB 14; Length 3365;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3365; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCGCGCGGCTGAGAGCCCTGTGGACAACTCGTCATTGTGAGGCACAGCGGTAGAC 60
Db |||||
QY 1 GCGCGCGGCTGAGAGCCCTGTGGACAACTCGTCATTGTGAGGCACAGCGGTAGAC 60
Db |||||
QY 61 CCTGCTTCTTAAGTGGGAGAACCTCTCCCTTCTCTGAGAAAGAAAGATGTCGAATGG 120
Db |||||
QY 61 CCTGCTTCTTAAGTGGGAGAACCTCTCCCTTCTCTGAGAAAGAAAGATGTCGAATGG 120
Db |||||
QY 121 CAGCACCATCTGCTTGGGAGAACCTCTCCCTTCTCTGAGAAAGAAAGATGTCGAATGG 180
Db |||||
QY 121 CAGCACCATCTGCTTGGGAGAACCTCTCCCTTCTCTGAGAAAGAAAGATGTCGAATGG 180
Db |||||
QY 181 TATCCACAGACGAGAAATTTCCGCTATCTCATCTCGTGTTCAGGGCCAGGGTGAAGAG 240
Db |||||
QY 181 TATCCACAGACGAGAAATTTCCGCTATCTCATCTCGTGTTCAGGGCCAGGGTGAAGAG 240
Db |||||
QY 241 TACATCCAGGTGAGCCCTGTGTGGACTACCTGACCTTCTGCTGCTGAGAGGTGAAGAG 300
Db |||||
QY 241 TACATCCAGGTGAGCCCTGTGTGGACTACCTGACCTTCTGCTGCTGAGAGGTGAAGAG 300
Db |||||
QY 301 CAGATTCAGAGGACAGTCGCCACCTCCGGGAACATGCAGGCAGTTGAACTGCTGCTGAGC 360
Db |||||
QY 301 CAGATTCAGAGGACAGTCGCCACCTCCGGGAACATGCAGGCAGTTGAACTGCTGCTGAGC 360
Db |||||
QY 361 ACCTTGGAGAGGGAGTCTGGCACTTGGTGGACTCGGGAATTCGTGGAGGCCCTCCGG 420
Db |||||
QY 361 ACCTTGGAGAGGGAGTCTGGCACTTGGTGGACTCGGGAATTCGTGGAGGCCCTCCGG 420
Db |||||
QY 421 AGAACCGGAGCCCTCTGGCCGCCGCTACATGAACCTTGAGCTCACGGACTTGCCTCT 480
Db |||||
QY 421 AGAACCGGAGCCCTCTGGCCGCCGCTACATGAACCTTGAGCTCACGGACTTGCCTCT 480
Db |||||
QY 481 CCATCGTTTGAGAACGCTCATGATGAATATCTCCAACCTGCTGAACCTTCTTCAGCCACT 540
Db |||||
QY 481 CCATCGTTTGAGAACGCTCATGATGAATATCTCCAACCTGCTGAACCTTCTTCAGCCACT 540
Db |||||
QY 541 CTGGTGGACAAAGCTTCTAGTTAGAGACGCTCTTGGATAAGTGCATGGAGGAACTGTTG 600
Db |||||
QY 541 CTGGTGGACAAAGCTTCTAGTTAGAGACGCTCTTGGATAAGTGCATGGAGGAACTGTTG 600
Db |||||

Db 2761 AAACCAGAGGATGCTCATAGATTTTGGAAATTACAGATGCAAGATATAATGGAAAG 2820

Qy 2821 AAAATGAAAAACCAAGAGAAATATTGCCAAGCATTA CAAGAATAACCCATCACTAATAACT 2880

Db 2821 AAAATGAAAAACCAAGAGAAATATTGCCAAGCATTA CAAGAATAACCCATCACTAATAACT 2880

Qy 2881 TTCCTTTGCAAAAACTGCAGTGTGCTAGCCTGTTCTGGGGAAGATATCCATGTAATTGAG 2940

Db 2881 TTCCTTTGCAAAAACTGCAGTGTGCTAGCCTGTTCTGGGGAAGATATCCATGTAATTGAG 2940

Qy 2941 AAAATGCATCACGTCAATATGACCCCGAATTCAGGAATTCAGGAACCTTTACATTTGAAGAGAAAC 3000

Db 2941 AAAATGCATCACGTCAATATGACCCCGAATTCAGGAATTCAGGAACCTTTACATTTGAAGAGAAAC 3000

Qy 3001 AAAGCACTGCAAAAGAAAGTGTGCCGACTATCAAAATAAATGGTGAATCATCTGCAAAATGT 3060

Db 3001 AAAGCACTGCAAAAGAAAGTGTGCCGACTATCAAAATAAATGGTGAATCATCTGCAAAATGT 3060

Qy 3061 GGCCAGGCTTGGGGAACAATGATGGTGCACAAAGGCTTAGATTTGCCTTGCTCAAAATA 3120

Db 3061 GGCCAGGCTTGGGGAACAATGATGGTGCACAAAGGCTTAGATTTGCCTTGCTCAAAATA 3120

Qy 3121 AGGAATTTTGTAGTGGTTTCAAAAATAATTCAAAAGAAACAATACAAAAAAGTGGGTA 3180

Db 3121 AGGAATTTTGTAGTGGTTTCAAAAATAATTCAAAAGAAACAATACAAAAAAGTGGGTA 3180

Qy 3181 GAATTACCTATCACATTTCCCAATCTTGACTATTCAGAAATGCTGTTATTAGTGATGAG 3240

Db 3181 GAATTACCTATCACATTTCCCAATCTTGACTATTCAGAAATGCTGTTATTAGTGATGAG 3240

Qy 3241 GATTAGCACTTGATTGAAGATTCTTTTAAATACTATCAGTTAAACATTTAATATGATTA 3300

Db 3241 GATTAGCACTTGATTGAAGATTCTTTTAAATACTATCAGTTAAACATTTAATATGATTA 3300

Qy 3301 TGATTAATGTATTCAATTATGCTACAGAACTGACATAAGAATCAATAAAATGATTGTTTTA 3360

Db 3301 TGATTAATGTATTCAATTATGCTACAGAACTGACATAAGAATCAATAAAATGATTGTTTTA 3360

Qy 3361 CTCTG 3365

Db 3361 CTCTG 3365

RESULT 3

US-10-755-889-405

; Sequence 405, Application US/10755889

; Publication No. US20040171823A1

; GENERAL INFORMATION:

; APPLICANT: Bristol-Myers Squibb Company

; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-kB

; TITLE OF INVENTION: PATHWAY

; FILE REFERENCE: D0284 NP

; CURRENT APPLICATION NUMBER: US/10/755,889

; CURRENT FILING DATE: 2004-01-13

; PRIOR APPLICATION NUMBER: U.S. 60/440,068

; PRIOR FILING DATE: 2003-01-14

; PRIOR APPLICATION NUMBER: U.S. 60/469,757

; PRIOR FILING DATE: 2003-05-12

; NUMBER OF SEQ ID NOS: 823

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 405

; LENGTH: 3380

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-755-889-405

Query Match 100.0%; Score 3365; DB 17; Length 3380;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 3365; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGCGCCGGCCTGAGAGCCCTGTGGACAACTCGTCATTGTCAGGCACAGAGCGGTAGAC 60

Db 1 GCGCGCCGGCCTGAGAGCCCTGTGGACAACTCGTCATTGTCAGGCACAGAGCGGTAGAC 60

Qy 61 CCTGCTTCTTAAGTGGGCAGCGGCACAGCGGCACGCACATTTACCTGTCCCGCAGACAA 120

Db 61 CCTGCTTCTTAAGTGGGCAGCGGCACAGCGGCACGCACATTTACCTGTCCCGCAGACAA 120

Qy 121 CAGCACCATCTGCTTGGGAGAACCCCTCTCCCTTCTCTGAGAAAAGAAAGATGTCGAATGGG 180

Db 121 CAGCACCATCTGCTTGGGAGAACCCCTCTCCCTTCTCTGAGAAAAGAAAGATGTCGAATGGG 180

Qy 181 TATTCACACAGACGAGAAATTTCCGCTATCTCATCTCGTCTCAGGGCCAGGGTGAATAATG 240

Db 181 TATTCACACAGACGAGAAATTTCCGCTATCTCATCTCGTCTCAGGGCCAGGGTGAATAATG 240

Qy 241 TACATCCAGTGGAGCCTGTGTGGAATCTGACTACCTGACCTTTCTGCTGCAGAGGTGAAGGAG 300

Db 241 TACATCCAGTGGAGCCTGTGTGGAATCTGACTACCTGACCTTTCTGCTGCAGAGGTGAAGGAG 300

Qy 301 CAGATTCAGAGACAGTCGCCACCTCCGGGAACATGCAGGCAAGTTGAACTGCTGCTGAGC 360

Db 301 CAGATTCAGAGACAGTCGCCACCTCCGGGAACATGCAGGCAAGTTGAACTGCTGCTGAGC 360

Qy 361 ACCTTGGAGAGGGAGTCTGGCACCTTGGTTGGACTCGGGGAATTCGTGGAGGCCCTCCGG 420

Db 361 ACCTTGGAGAGGGAGTCTGGCACCTTGGTTGGACTCGGGGAATTCGTGGAGGCCCTCCGG 420

Qy 421 AGAACCGGCAGCCCTCTGGCCGCCCTACATGAACCCCTGAGTCA CGGACTTGCCCTCT 480

Db 421 AGAACCGGCAGCCCTCTGGCCGCCCTACATGAACCCCTGAGTCA CGGACTTGCCCTCT 480

Qy 481 CCATCGTTTGGAGAACGCTCATGATGAATATCTCCAACCTGCTGAACCTCTTCAGCCCACT 540

Db 481 CCATCGTTTGGAGAACGCTCATGATGAATATCTCCAACCTGCTGAACCTCTTCAGCCCACT 540

Qy 541 CTGGTGGACAAGCTTCTAGTTAGAGACGCTTTGGATAAGTGCATGGAGGAGAACTGTTG 600

Db 541 CTGGTGGACAAGCTTCTAGTTAGAGACGCTTTGGATAAGTGCATGGAGGAGAACTGTTG 600

Qy 601 ACAATTGAAGACAGAAAAACCGGATTGCTGCTGCAGAAAAACAATGGAATGAATCAGGTGTA 660

Db 601 ACAATTGAAGACAGAAAAACCGGATTGCTGCTGCAGAAAAACAATGGAATGAATCAGGTGTA 660

Qy 661 AGAGAGCTACTAAAAAGGATTGTGCAGAAAAAGAAAACTGGTTCTCTGCATTTCTGAATGTT 720

Db 661 AGAGAGCTACTAAAAAGGATTGTGCAGAAAAAGAAAACTGGTTCTCTGCATTTCTGAATGTT 720

Qy 721 CTTCTGTCAAAACAGGAAAAACAATGAACCTGTCCAAGATTTAACAGGCTCTGATTGCTCAGAA 780

Db 721 CTTCTGTCAAAACAGGAAAAACAATGAACCTGTCCAAGATTTAACAGGCTCTGATTGCTCAGAA 780

Qy 781 AGCAATGCAGAGATTGAGAAATTTATCACAAGTTGATGGTCTCTCAAGTGGAAAGAGCAACTT 840

Db 781 AGCAATGCAGAGATTGAGAAATTTATCACAAGTTGATGGTCTCTCAAGTGGAAAGAGCAACTT 840

Qy 841 CTTTCAACCACAGTTTCAGCCCAAATCTGGAGAAAGGAGGTCTGGGCGCATGGAGAAATACTCA 900

Db 841 CTTTCAACCACAGTTTCAGCCCAAATCTGGAGAAAGGAGGTCTGGGCGCATGGAGAAATACTCA 900

Qy 901 TCAGAAATCATCTTTTGCAGATTCTTCTGTAGTTTCAGAAATCAGACACAAGTTTGGCAGAA 960

Db 901 TCAGAAATCATCTTTTGCAGATTCTTCTGTAGTTTCAGAAATCAGACACAAGTTTGGCAGAA 960

Qy 961 GGAAGTGTACAGTCTTAGATGAAGTCTTGGACATAACAGCAACATGGGCAGTGATTCA 1020

Db 961 GGAAGTGTACAGTCTTAGATGAAGTCTTGGACATAACAGCAACATGGGCAGTGATTCA 1020

Qy 1021 GGCACCATGGGAAGTGATTTCAGATGAAGAGAAATGTGGCAGCAAGAGCATCCCCGGAGCCA 1080

Db 1021 GGCACCATGGGAAGTGATTTCAGATGAAGAGAAATGTGGCAGCAAGAGCATCCCCGGAGCCA 1080

Qy 1081 GAACTCCAGCTCAGGCCCTTACCATAATGGAAGTTGCCAGCCAGCCCTTGGAAAGGAAGAT 1140

Db 1081 GAACTCCAGCTCAGGCCCTTACCATAATGGAAGTTGCCAGCCAGCCCTTGGAAAGGAAGAT 1140

QY 1141 ATCATCATCTGCCTCCCTACAGGAGTGGAAAAACACAGAGTGGCTGTTTACATTGCCAAG 1200
Db 1141 ATCATCATCTGCCTCCCTACAGGAGTGGAAAAACACAGAGTGGCTGTTTACATTGCCAAG 1200
QY 1201 GATCATTAGACAAGAAAGAAAAAGCATCTGAGCCTGGAAAAAGTTATAGTTCTTGTCAAT 1260
Db 1201 GATCATTAGACAAGAAAGAAAAAGCATCTGAGCCTGGAAAAAGTTATAGTTCTTGTCAAT 1260
QY 1261 AAGGTACTGTAGTTGAACAGCTCTTCCGCAAGGAGTTCCAAACCATTTTGAAGAAATGG 1320
Db 1261 AAGGTACTGTAGTTGAACAGCTCTTCCGCAAGGAGTTCCAAACCATTTTGAAGAAATGG 1320
QY 1321 TATCGTGTATTGGATTAAAGTGGTGATACCCAACTGAAAAATATCATTTCCAGAGTTGTC 1380
Db 1321 TATCGTGTATTGGATTAAAGTGGTGATACCCAACTGAAAAATATCATTTCCAGAGTTGTC 1380
QY 1381 AAGTCCTGTGATATTATTATCAGTACAGCTCAAAATCCTTGAAAACTCCCTCTTAAACTTG 1440
Db 1381 AAGTCCTGTGATATTATTATCAGTACAGCTCAAAATCCTTGAAAACTCCCTCTTAAACTTG 1440
QY 1441 GAAAAATGGAGAAGATGCTGGTGTCAATTTGTGAGACTTTTCCCTCATTTATCATGATGAA 1500
Db 1441 GAAAAATGGAGAAGATGCTGGTGTCAATTTGTGAGACTTTTCCCTCATTTATCATGATGAA 1500
QY 1501 TGTCATCACACCAACAAAGAAAGCAGTGTATAATAACATCATGAGGCATTTATTTGATGCAG 1560
Db 1501 TGTCATCACACCAACAAAGAAAGCAGTGTATAATAACATCATGAGGCATTTATTTGATGCAG 1560
QY 1561 AAGTTGAAAAACAAATAGACTCAAGAAAGAAAAACAAACACAGTATTTCCCTCCCTCAGATA 1620
Db 1561 AAGTTGAAAAACAAATAGACTCAAGAAAGAAAAACAAACACAGTATTTCCCTCCCTCAGATA 1620
QY 1621 CTGGGACTAACAGCTTACCTGGTGTGGAGGGCCACGAAAGCAAGCCAAAGCTGAAGAA 1680
Db 1621 CTGGGACTAACAGCTTACCTGGTGTGGAGGGCCACGAAAGCAAGCCAAAGCTGAAGAA 1680
QY 1681 CACATTTTAAAACTATGTGCCAATCTTGATGCATTTTACTATTAAAACTGTTAAAGAAAAAC 1740
Db 1681 CACATTTTAAAACTATGTGCCAATCTTGATGCATTTTACTATTAAAACTGTTAAAGAAAAAC 1740
QY 1741 CTTGATCAACTGAAAAACCAAATACAGGAGCCATGCAAGAAAGTTTGCCATTGCGAGATGCA 1800
Db 1741 CTTGATCAACTGAAAAACCAAATACAGGAGCCATGCAAGAAAGTTTGCCATTGCGAGATGCA 1800
QY 1801 ACCAGAGAAGATCCATTTAAAGAGAAACTTCTAGAAATATGACAAGGATTCAAACTTAT 1860
Db 1801 ACCAGAGAAGATCCATTTAAAGAGAAACTTCTAGAAATATATGACAAGGATTCAAACTTAT 1860
QY 1861 TGTCAAATGAGTCCAATGTGAGATTTTGGAACTCAACCCCTATGAAACAAATGGGCCATTCAA 1920
Db 1861 TGTCAAATGAGTCCAATGTGAGATTTTGGAACTCAACCCCTATGAAACAAATGGGCCATTCAA 1920
QY 1921 ATGGAAAAAAAGCTGCAAAAAAAGGAAATCGCAAAAGAACGTTTGTGCGAGAACATTTG 1980
Db 1921 ATGGAAAAAAAGCTGCAAAAAAAGGAAATCGCAAAAGAACGTTTGTGCGAGAACATTTG 1980
QY 1981 AGGAAGTACAATGAGGCCCTACAAATTAATGACACAAATTCGAATGATAGATGCGTATACT 2040
Db 1981 AGGAAGTACAATGAGGCCCTACAAATTAATGACACAAATTCGAATGATAGATGCGTATACT 2040
QY 2041 CATCTTGAAACTTTCTATAATGAAGAGAAAGATAAGAAAGTTTGCAGTCATAGAAGATGAT 2100
Db 2041 CATCTTGAAACTTTCTATAATGAAGAGAAAGATAAGAAAGTTTGCAGTCATAGAAGATGAT 2100
QY 2101 AGTGATGAGGGTGGTGATGATGAGTATTGTGATGGTGATGAAGATGAGGATGATTTAAAG 2160
Db 2101 AGTGATGAGGGTGGTGATGATGAGTATTGTGATGGTGATGAAGATGAGGATGATTTAAAG 2160
QY 2161 AAACCTTTGAAACTGGATGAAACAGATAGATTTCTCATGACTTTATTTTTTGAACAAAT 2220
Db 2161 AAACCTTTGAAACTGGATGAAACAGATAGATTTCTCATGACTTTATTTTTTGAACAAAT 2220
QY 2221 AAAATGTTGAAAGGCTGGCTGAAAAACCCAGAAATATGAAAAATGAAAAAGCTGACCAAAATTA 2280

Db 2221 AAAATGTTGAAAGGCTGGCTGAAAAACCCAGAAATATGAAAAATGAAAAAGCTGACCAAAATTA 2280
QY 2281 AGAAATACCATAATGGAGCAATATACTAGGACTGAGGAATCAGCACGAGGAATAATCTTT 2340
Db 2281 AGAAATACCATAATGGAGCAATATACTAGGACTGAGGAATCAGCACGAGGAATAATCTTT 2340
QY 2341 ACAAAAAACACGACAGAGTGCATATGCGCTTTCCAGTGGATTACTGAAAAATGAAAAATTT 2400
Db 2341 ACAAAAAACACGACAGAGTGCATATGCGCTTTCCAGTGGATTACTGAAAAATGAAAAATTT 2400
QY 2401 GCTGAAGTAGGAGTCAAAAGCCCCACCATCTGATTGGAGCTGGACACAGCAGTGAGTTCAAA 2460
Db 2401 GCTGAAGTAGGAGTCAAAAGCCCCACCATCTGATTGGAGCTGGACACAGCAGTGAGTTCAAA 2460
QY 2461 CCCATGACACAGAAATGAAACAAAAAGAGTCAATTAGTAAATTTTCGCACTGGAAAAATCAAT 2520
Db 2461 CCCATGACACAGAAATGAAACAAAAAGAGTCAATTAGTAAATTTTCGCACTGGAAAAATCAAT 2520
QY 2521 CTGCTTATCGCTACCAAGTGGCAGAAAGTCTGGATATTTAAAGAAATGTAACATTGTT 2580
Db 2521 CTGCTTATCGCTACCAAGTGGCAGAAAGTCTGGATATTTAAAGAAATGTAACATTGTT 2580
QY 2581 ATCCGTTATGGTCTCGTCACCAATGAAATAGCCATGGTCCAGGCCCGTGGTCCAGCCAGA 2640
Db 2581 ATCCGTTATGGTCTCGTCACCAATGAAATAGCCATGGTCCAGGCCCGTGGTCCAGCCAGA 2640
QY 2641 GCTGATGAGAGCACCTACGTCTGTTGCTCAGAGTGGTTCAGGAGTTATCGAACATGAG 2700
Db 2641 GCTGATGAGAGCACCTACGTCTGTTGCTCAGAGTGGTTCAGGAGTTATCGAACATGAG 2700
QY 2701 ACAGTTAATGATTTCCGAGAGAAGATGATGTATAAAGCTATACATTGTTGTTCAAAATATG 2760
Db 2701 ACAGTTAATGATTTCCGAGAGAAGATGATGTATAAAGCTATACATTGTTGTTCAAAATATG 2760
QY 2761 AAACAGAGGAGTATGCTCATAAAGATTTTGGAAATACAGATGCAAAAGTATAATGGAAG 2820
Db 2761 AAACAGAGGAGTATGCTCATAAAGATTTTGGAAATACAGATGCAAAAGTATAATGGAAG 2820
QY 2821 AAAATGAAAAACCAAGAGAAATATTGCAAGCATTTACAAGAAATACCCATCACTAATACT 2880
Db 2821 AAAATGAAAAACCAAGAGAAATATTGCAAGCATTTACAAGAAATACCCATCACTAATACT 2880
QY 2881 TTCCCTTTGCAAAAACTGCAAGTGTGCTAGCCTGTTCTGGGGAAGATATCCATGTAATTGAG 2940
Db 2881 TTCCCTTTGCAAAAACTGCAAGTGTGCTAGCCTGTTCTGGGGAAGATATCCATGTAATTGAG 2940
QY 2941 AAAATGCAATCAGTCAATATGACCCAGAAATTCAGAGAACTTTTACATTGTAAGAGAAAAAC 3000
Db 2941 AAAATGCAATCAGTCAATATGACCCAGAAATTCAGAGAACTTTTACATTGTAAGAGAAAAAC 3000
QY 3001 AAAGCACTGCAAAAGAAAGTGTGCCGACTATCAAAATAAATGGTGAATCATCTGCAAAATGT 3060
Db 3001 AAAGCACTGCAAAAGAAAGTGTGCCGACTATCAAAATAAATGGTGAATCATCTGCAAAATGT 3060
QY 3061 GGCCAGGCTTGGGGAACAATGATGGTGACAAAGGCTTAGATTGCTCTCAAAATA 3120
Db 3061 GGCCAGGCTTGGGGAACAATGATGGTGACAAAGGCTTAGATTGCTCTCAAAATA 3120
QY 3121 AGGAATTTGTAGTGGTTTCAAAAAATAATTCAACAAAGAAACAATACAAAAAGTGGGTA 3180
Db 3121 AGGAATTTGTAGTGGTTTCAAAAAATAATTCAACAAAGAAACAATACAAAAAGTGGGTA 3180
QY 3181 GAATTACCTATCACATTTCCCAATCTTGACTATTAGAAATGCTGTTTATTATTAGTGATGAG 3240
Db 3181 GAATTACCTATCACATTTCCCAATCTTGACTATTAGAAATGCTGTTTATTATTAGTGATGAG 3240
QY 3241 GATTAGCACTTGATTGAAGATCTTTTAAATACATCAAGTTAAACATTTTAAATGATTA 3300
Db 3241 GATTAGCACTTGATTGAAGATCTTTTAAATACATCAAGTTAAACATTTTAAATGATTA 3300
QY 3301 TGATTAAATGATTTCAATTATGCTACAGAACTGACATAAGAAATCAATAAAATGATGTTTTTA 3360

```
Db 3301 TGATTAATGTATTCAATTATGCTACAGAACTGACATAAGAAATCAATAAAATGATTGTTTTA 3360
QY 3361 CTCTG 3365
Db 3361 CTCTG 3365

RESULT 4
US-10-055-475-8
; Sequence 8, Application US/10055475
; Publication No. US20030022855A1
; GENERAL INFORMATION:
; APPLICANT: Fisher, Paul B.
; APPLICANT: Kang, Dong-Chul
; APPLICANT: Gopalkrishnan, Rahul V.
; TITLE OF INVENTION: USE OF MDA-5 AS AN ANTIVIRAL AND
; TITLE OF INVENTION: ANTIPROLIFERATIVE AGENT
; FILE REFERENCE: A34614-A-PCT-USA-A (070050.1921)
; CURRENT APPLICATION NUMBER: US/10/055,475
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: PCT/US01/06960
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 09/515,363
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 3365
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-055-475-8

Query Match 99.9%; Score 3361.8; DB 14; Length 3365;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3363; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GCGGCGCGCCTGAGAGCCCTGTGGACAACCTCGTCATTGTGAGGCACAGAGCGGTAGAC 60
Db 1 GCGGCGCGCCTGAGAGCCCTGTGGACAACCTCGTCATTGTGAGGCACAGAGCGGTAGAC 60

QY 61 CCTGCTTCTTAAGTGGGCAGCGGACAGCGGCACGCACATTTCACCTGTCCCGCAGACAA 120
Db 61 CCTGCTTCTTAAGTGGGCAGCGGACAGCGGCACGCACATTTCACCTGTCCCGCAGACAA 120

QY 121 CAGCACCATCTGCTTGGAGAACCTCTCCCTTCTCTGAGAAAGAAAGATGTCGAATGGG 180
Db 121 CAGCACCATCTGCTTGGAGAACCTCTCCCTTCTCTGAGAAAGAAAGATGTCGAATGGG 180

QY 181 TATTCACAGACGAGAAATTTCCGCTATCTCATCTCGTGCITTCAGGGCCAGGGTGAAATG 240
Db 181 TATTCACAGACGAGAAATTTCCGCTATCTCATCTCGTGCITTCAGGGCCAGGGTGAAATG 240

QY 241 TACATCCAGGTGAGACCTGTGCTGGACTACCTGACCTTTCTGCCTGCAGAGGTGAAGGAG 300
Db 241 TACATCCAGGTGAGACCTGTGCTGGACTACCTGACCTTTCTGCCTGCAGAGGTGAAGGAG 300

QY 301 CAGATTCAGAGGACAGTCGCCACCTCCGGGAACATGCAGGCACTTGAACCTGCTGCTGAGC 360
Db 301 CAGATTCAGAGGACAGTCGCCACCTCCGGGAACATGCAGGCACTTGAACCTGCTGCTGAGC 360

QY 361 ACCTTGGAGAGGAGTCTGGCACCTTGGTTGGACTCGGGAATTCGTGGAGGCCCTCCGG 420
Db 361 ACCTTGGAGAGGAGTCTGGCACCTTGGTTGGACTCGGGAATTCGTGGAGGCCCTCCGG 420

QY 421 AGAACCGGCAGCCCTCTGGCCGCCCGCTACATGAACCCCTGAGCTCACGGACTTGCCTCT 480
Db 421 AGAACCGGCAGCCCTCTGGCCGCCCGCTACATGAACCCCTGAGCTCACGGACTTGCCTCT 480

QY 481 CCATCGTTTGAACCGCTCATGATGAATATCTCCAACTGCTGAACCTCTTCAGCCCCACT 540
Db 481 CCATCGTTTGAACCGCTCATGATGAATATCTCCAACTGCTGAACCTCTTCAGCCCCACT 540

QY 541 CTGGTGGACAAGCTTCTAGTTAGAGACGCTCTTGATTAAGTGCAATGGAGGAACTGTTG 600
```

```
Db 541 CTGGTGGACAAGCTTCTAGTTAGAGACGTCTTGGATAAGTGTGATGGAGGAACTGTTG 600
QY 601 ACAATTGAAGACAGAAACCGGATTGCTGCTGCAGAAAAACAATGGAATGAATCAGGTGTA 660
Db 601 ACAATTGAAGACAGAAACCGGATTGCTGCTGCAGAAAAACAATGGAATGAATCAGGTGTA 660
QY 661 AGAGAGCTACTAAAAAGGATTGTGCAGAAAAAACAATGTTCTCTGCAATTTCTGAATGTT 720
Db 661 AGAGAGCTACTAAAAAGGATTGTGCAGAAAAAACAATGTTCTCTGCAATTTCTGAATGTT 720
QY 721 CTTTCGTCAAACAGGAAACAATGAATTTTATCACAAGTTTCAAGAGTTAACAGGCTCTGATTGCTCAGAA 780
Db 721 CTTTCGTCAAACAGGAAACAATGAATTTTCAAGAGTTAACAGGCTCTGATTGCTCAGAA 780
QY 781 AGCAATGCAGAGATTGAGAAATTTATCACAAGTTTATGATGGTCTCAAGTGGAGAGCAACTT 840
Db 781 AGCAATGCAGAGATTGAGAAATTTATCACAAGTTTATGATGGTCTCAAGTGGAGAGCAACTT 840
QY 841 CTTTCAACCACAGTTTCAGCCAAATCTGGAGAAAGGAGTCTGGGCGATGGAGAAATACTCA 900
Db 841 CTTTCAACCACAGTTTCAGCCAAATCTGGAGAAAGGAGTCTGGGCGATGGAGAAATACTCA 900
QY 901 TCAGAAATCATCTTTTGCAGATTCTTCTGTAGTTTTCAGAAATCAGACACAAAGTTTGGCAGAA 960
Db 901 TCAGAAATCATCTTTTGCAGATTCTTCTGTAGTTTTCAGAAATCAGACACAAAGTTTGGCAGAA 960
QY 961 GGAAGTGTGAGCTGCTTAGATGAAGTCTTGGACATAACAGCAACATGGGCAGTGATTCA 1020
Db 961 GGAAGTGTGAGCTGCTTAGATGAAGTCTTGGACATAACAGCAACATGGGCAGTGATTCA 1020
QY 1021 GGCACCATGGGAAGTGATTTCAGATGAAGAGAAATGTGGCAGCAAGAGCATCCCCGGAGCCA 1080
Db 1021 GGCACCATGGGAAGTGATTTCAGATGAAGAGAAATGTGGCAGCAAGAGCATCCCCGGAGCCA 1080
QY 1081 GAACCTCCAGCTCAGGCCCTTACCAATGGAAGTTGCCAGCCAGCCCTTGGAAAGGGAAGAT 1140
Db 1081 GAACCTCCAGCTCAGGCCCTTACCAATGGAAGTTGCCAGCCAGCCCTTGGAAAGGGAAGAT 1140
QY 1141 ATCATCATCTGCCTCCCTACAGGGAGTGGAAAAACCAGAGTGGCTGTTTACATTGCCAAG 1200
Db 1141 ATCATCATCTGCCTCCCTACAGGGAGTGGAAAAACCAGAGTGGCTGTTTACATTGCCAAG 1200
QY 1201 GATCATTAGACAAGAAAGAAAAAGCATCTGAGCCTGGAAAGTTATAGTTCTTGTCAAT 1260
Db 1201 GATCATTAGACAAGAAAGAAAAAGCATCTGAGCCTGGAAAGTTATAGTTCTTGTCAAT 1260
QY 1261 AAGGTACTGCTAGTTGAACAGCTCTTCCGCAAGGAGTTCCAACCATTTTGAAGAAATGG 1320
Db 1261 AAGGTACTGCTAGTTGAACAGCTCTTCCGCAAGGAGTTCCAACCATTTTGAAGAAATGG 1320
QY 1321 TATCGTGTATTGGATTAAAGTGGTGATACCCAACTGAAATATATCATTTCCAGAAGTTGTC 1380
Db 1321 TATCGTGTATTGGATTAAAGTGGTGATACCCAACTGAAATATATCATTTCCAGAAGTTGTC 1380
QY 1381 AAGTCCTGTGATATTATTATCAGTACAGCTCAAAATCCCTGAAAACTCCCTCTTAAACTTG 1440
Db 1381 AAGTCCTGTGATATTATTATCAGTACAGCTCAAAATCCCTGAAAACTCCCTCTTAAACTTG 1440
QY 1441 GAAAAATGGAGAAGATGCTGGTGTCAATTGTGAGACTTTTCCCTCATTTATCATTGATGAA 1500
Db 1441 GAAAAATGGAGAAGATGCTGGTGTCAATTGTGAGACTTTTCCCTCATTTATCATTGATGAA 1500
QY 1501 TGTATCATCACCAACAAGAAAGAGTGATATAATAACATCATGAGGCATTATTTGATGAG 1560
Db 1501 TGTATCATCACCAACAAGAAAGAGTGATATAATAACATCATGAGGCATTATTTGATGAG 1560
QY 1561 AAGTTGAAAAACAATAGACTCAAGAAAGAAAAACAACCCAGTGATTCCCTTCCCTCAGATA 1620
Db 1561 AAGTTGAAAAACAATAGACTCAAGAAAGAAAAACAACCCAGTGATTCCCTTCCCTCAGATA 1620
QY 1621 CTGGGACTAACAGCTTTCACCTGGTGTGGAGGGGCCACGAAGCAAGCCAAAGCTGAAGAA 1680
Db 1621 CTGGGACTAACAGCTTTCACCTGGTGTGGAGGGGCCACGAAGCAAGCCAAAGCTGAAGAA 1680
```

Db 1621 CTGGGACTAACAGCTTCACCTGGTGTGGAGGGGCCACGAAGCAAGCCAAAGCTGAAGAA 1680

QY 1681 CACATTTTAAACACTATGTGCCAATCTTGATGCATTTACTATTAAACCTGTTAAAGAAAAC 1740

Db 1681 CACATTTTAAACACTATGTGCCAATCTTGATGCATTTACTATTAAACCTGTTAAAGAAAAC 1740

QY 1741 CTTGATCAACTGAAAAACCAAATACAGGAGCCATGCAAGAAAGTTTGCCATTGCAGATGCA 1800

Db 1741 CTTGATCAACTGAAAAACCAAATACAGGAGCCATGCAAGAAAGTTTGCCATTGCAGATGCA 1800

QY 1801 ACCAGAGAGATCCATTTAAAGAGAAAACCTTCTAGAAATAATGACAAGGATTCAAAACCTTAT 1860

Db 1801 ACCAGAGAGATCCATTTAAAGAGAAAACCTTCTAGAAATAATGACAAGGATTCAAAACCTTAT 1860

QY 1861 TGTCAAATGAGTCCAAATGTCAGATTTTGGAACTCAACCTATGAACAATGGGCCATTCAA 1920

Db 1861 TGTCAAATGAGTCCAAATGTCAGATTTTGGAACTCAACCTATGAACAATGGGCCATTCAA 1920

QY 1921 ATGGAAAAAAAGCTGCAAAAAAAGGAAATCGCAAAAGAACGTGTTGTGCGAACATTGTG 1980

Db 1921 ATGGAAAAAAAGCTGCAAAAGAGGAAATCGCAAAAGAACGTGTTGTGCGAACATTGTG 1980

QY 1981 AGGAAGTACAATGAGGCCCTACAAATTAATGACACAAATTCGAATGATAGATGCGTATACT 2040

Db 1981 AGGAAGTACAATGAGGCCCTACAAATTAATGACACAAATTCGAATGATAGATGCGTATACT 2040

QY 2041 CATCTTGAAACCTTCTATATAATGAAGAGAAAAGATAAGAGTTTGCAGTCATAGAGATGAT 2100

Db 2041 CATCTTGAAACCTTCTATATAATGAAGAGAAAAGATAAGAGTTTGCAGTCATAGAGATGAT 2100

QY 2101 AGTGATGAGGGTGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2160

Db 2101 AGTGATGAGGGTGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2160

QY 2161 AAACCTTTGAAACTGGATGAAACAGATAGATTTCTCATGACTTTATTTTTGAAAAACAAT 2220

Db 2161 AAACCTTTGAAACTGGATGAAACAGATAGATTTCTCATGACTTTATTTTTGAAAAACAAT 2220

QY 2221 AAAATGTTGAAAAGGCTGGCTGAAAAACCCAGAAATATGAAAATGAAAAGCTGACCAATTA 2280

Db 2221 AAAATGTTGAAAAGGCTGGCTGAAAACCCAGAAATATGAAAATGAAAAGCTGACCAATTA 2280

QY 2281 AGAAATACCAATAATGGAGCAATATCTAGGACTGAGGAATCAGCAGGAGGAATATCTTT 2340

Db 2281 AGAAATACCAATAATGGAGCAATATCTAGGACTGAGGAATCAGCAGGAGGAATATCTTT 2340

QY 2341 ACAAACACGACAGAGTGCCATATCGCTTTCCAGTGGATTACTGAAAAATGAAAAATTT 2400

Db 2341 ACAAACACGACAGAGTGCCATATCGCTTTCCAGTGGATTACTGAAAAATGAAAAATTT 2400

QY 2401 GCTGAAGTAGGAGTCAAAAGCCACCACCTCTGATTGGAGCTGGACACAGCAGTGAGTTCAA 2460

Db 2401 GCTGAAGTAGGAGTCAAAAGCCACCACCTCTGATTGGAGCTGGACACAGCAGTGAGTTCAA 2460

QY 2461 CCCATGACACAGAAATGAACAAAAAGAAAGTCAATTAGTAAATTTCCGACTGGAAAAATAAAT 2520

Db 2461 CCCATGACACAGAAATGAACAAAAAGAAAGTCAATTAGTAAATTTCCGACTGGAAAAATAAAT 2520

QY 2521 CTGCTTATCGTACCACAGTGGCAGAAAGAGTCTGGATATTAAGAAATGTAACATTGTT 2580

Db 2521 CTGCTTATCGTACCACAGTGGCAGAAAGAGTCTGGATATTAAGAAATGTAACATTGTT 2580

QY 2581 ATCCGTTATGGTCTCGTCACCAATGAAATAGCCATGTTCCAGGCCGCTGTCGAGCCAGA 2640

Db 2581 ATCCGTTATGGTCTCGTCACCAATGAAATAGCCATGTTCCAGGCCGCTGTCGAGCCAGA 2640

QY 2641 GCTGATGAGAGCACCTACGTCCTGGTGTCTCAGAGTGGTTCAGGAGTTATCGAACATGAG 2700

Db 2641 GCTGATGAGAGCACCTACGTCCTGGTGTCTCAGAGTGGTTCAGGAGTTATCGAACATGAG 2700

QY 2701 ACAGTTAATGATTTCCGAGAGAGATGATGTATAAGCTATACATTGTGTTCAAAATATG 2760

Db 2701 ACAGTTAATGATTTCCGAGAGAGATGATGTATAAGCTATACATTGTGTTCAAAATATG 2760

QY 2761 AAACCAGAGGAGTATGCTCATAAGATTTTGGAAATACAGATGCAAGATATAATGGAAG 2820

Db 2761 AAACCAGAGGAGTATGCTCATAAGATTTTGGAAATACAGATGCAAGATATAATGGAAG 2820

QY 2821 AAAATGAAACCAAGAGAAATATTGCCAAGCAATACAAAGAAATACCACACTACTAATACT 2880

Db 2821 AAAATGAAACCAAGAGAAATATTGCCAAGCAATACAAAGAAATACCACACTACTAATACT 2880

QY 2881 TTCTTTGCAAAAACTGCAGTGTGCTAGCCCTGTTCTGGGGAAGATATCCATGTAATTGAG 2940

Db 2881 TTCTTTGCAAAAACTGCAGTGTGCTAGCCCTGTTCTGGGGAAGATATCCATGTAATTGAG 2940

QY 2941 AAAATGCATCACGTCAATATGACCCCGAGAATTCAGGAACTTTTACATTGTAAGAGAAAAC 3000

Db 2941 AAAATGCATCACGTCAATATGACCCCGAGAATTCAGGAACTTTTACATTGTAAGAGAAAAC 3000

QY 3001 AAAGCACTGCAAAAAGAGTGTGCCGACTATCAAAATAAATGGTGAATCATCTGCAAAATGT 3060

Db 3001 AAAGCACTGCAAAAAGAGTGTGCCGACTATCAAAATAAATGGTGAATCATCTGCAAAATGT 3060

QY 3061 GGCCAGGCTTGGGGAACAATGATGGTGCACAAAGGCTTAGATTTGCCCTTGTCTCAAAATA 3120

Db 3061 GGCCAGGCTTGGGGAACAATGATGGTGCACAAAGGCTTAGATTTGCCCTTGTCTCAAAATA 3120

QY 3121 AGGAATTTTGTAGTGGTTTTTCAAAAATAAATTTCAACAAAGAAACAATACAAAAAGTGGGTA 3180

Db 3121 AGGAATTTTGTAGTGGTTTTTCAAAAATAAATTTCAACAAAGAAACAATACAAAAAGTGGGTA 3180

QY 3181 GAATTACCTATCACATTTCCCAATCTTGACTATTTCAGAAATGCTGTTTATTATTAGTGATGAG 3240

Db 3181 GAATTACCTATCACATTTCCCAATCTTGACTATTTCAGAAATGCTGTTTATTATTAGTGATGAG 3240

QY 3241 GATTAGCACCTTGATTGAAGATTTCTTTAAAAATACTATCAGTTAAACATTTTAATAATGATTA 3300

Db 3241 GATTAGCACCTTGATTGAAGATTTCTTTAAAAATACTATCAGTTAAACATTTTAATAATGATTA 3300

QY 3301 TGATTAATGTATTCTATGCTACAGAACTGACATAAGAAATCAATAAAATGATTGTTTAA 3360

Db 3301 TGATTAATGTATTCTATGCTACAGAACTGACATAAGAAATCAATAAAATGATTGTTTAA 3360

QY 3361 CTCTG 3365

Db 3361 CTCTG 3365

RESULT 5

US-10-198-846-13042

; Sequence 13042, Application US/10198846

; Publication No. US20030099974A1

; GENERAL INFORMATION:

; APPLICANT: Lillie, James

; APPLICANT: Xu, Yongyao

; APPLICANT: Wang, Youzhen

; APPLICANT: Steinmann, Kathleen

; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS

; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND

; TITLE OF INVENTION: THERAPY OF BREAST CANCER

; FILE REFERENCE: MRI-049

; CURRENT APPLICATION NUMBER: US/10/198,846

; CURRENT FILING DATE: 2002-07-18

; PRIOR APPLICATION NUMBER: 60/306,220

; PRIOR FILING DATE: 2001-07-18

; NUMBER OF SEQ ID NOS: 14084

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 13042

; LENGTH: 3928

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: 1, 3925, 3926, 3927, 3928

; OTHER INFORMATION: n = A,T,C or G

US-10-198-846-13042

Query Match		99.9%;	Score 3360.2;	DB 14;	Length 3928;		
Best Local Similarity		99.9%;	Pred. No. 0;				
Matches 3362;		Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;		
QY	1	GCGCGCGGCTGAGAGCCCTGTGGACAACCTCGTCATTTGTACGCACAGAGCGGTAGAC	60				
DB	443	GCGCGCGGCTGAGAGCCCTGTGGACAACCTCGTCATTTGTACGCACAGAGCGGTAGAC	502				
QY	61	CCTGCTTCTTAAGTGGGAGCGGACAGGCGCACGCACATTTCACTGTCCCGCAGACAA	120				
DB	503	CCTGCTTCTTAAGTGGGAGCGGACAGGCGCACGCACATTTCACTGTCCCGCAGACAA	562				
QY	121	CAGCACCATCTGCTTGGGAGAACCTCTCCCTTCTCTGAGAAAGAAAGATGTCGAATGG	180				
DB	563	CAGCACCATCTGCTTGGGAGAACCTCTCCCTTCTCTGAGAAAGAAAGATGTCGAATGG	622				
QY	181	TATTCCACAGACGAGAAATTTCCGCTATCTCATCTCGTCTTACGGGCCAGGGTGAATA	240				
DB	623	TATTCCACAGACGAGAAATTTCCGCTATCTCATCTCGTCTTACGGGCCAGGGTGAATA	682				
QY	241	TACATCCAGGTGGAGCTGTCTGGACTACCTGACCTTTCTGCCTGCAGAGGTGAAGGAG	300				
DB	683	TACATCCAGGTGGAGCTGTCTGGACTACCTGACCTTTCTGCCTGCAGAGGTGAAGGAG	742				
QY	301	CAGATTCAGAGGACAGTCGCCACCTCCGGGAAACATGCAGGCAGTTGAACTGCTGCTGAGC	360				
DB	743	CAGATTCAGAGGACAGTCGCCACCTCCGGGAAACATGCAGGCAGTTGAACTGCTGCTGAGC	802				
QY	361	ACCTTGGAGAAGGAGTCTGGCACCTTGGTTGGAATCTCGGGAAATTCGTGGAGGCCCTCCGG	420				
DB	803	ACCTTGGAGAAGGAGTCTGGCACCTTGGTTGGAATCTCGGGAAATTCGTGGAGGCCCTCCGG	862				
QY	421	AGAACCGGCAGCCCTCTGGCCGCCCGCTACATGAACCTGAGCTCACGGACTTGCCCTCT	480				
DB	863	AGAACCGGCAGCCCTCTGGCCGCCCGCTACATGAACCTGAGCTCACGGACTTGCCCTCT	922				
QY	481	CCATCGTTTGAGAACGCTCATGATGAATATCTCCAACTGCTGAACCTCCTTCAGCCCCACT	540				
DB	923	CCATCGTTTGAGAACGCTCATGATGAATATCTCCAACTGCTGAACCTCCTTCAGCCCCACT	982				
QY	541	CTGGTGGACAAGCTTCTAGTTAGAGACGTCCTTGATAAGTGCATGGAGGAGGAAC	600				
DB	983	CTGGTGGACAAGCTTCTAGTTAGAGACGTCCTTGATAAGTGCATGGAGGAGGAAC	1042				
QY	601	ACAATTGAAGACAGAAACCGGATTGCTGCTGCAGAAAAACAATGGAATGAATCAGGTGTA	660				
DB	1043	ACAATTGAAGACAGAAACCGGATTGCTGCTGCAGAAAAACAATGGAATGAATCAGGTGTA	1102				
QY	661	AGAGAGCTACTAAAAAGGATTGTGCAGAAAGAAAACTGGTTCTCTGCATTTCTGAATGTT	720				
DB	1103	AGAGAGCTACTAAAAAGGATTGTGCAGAAAGAAAACTGGTTCTCTGCATTTCTGAATGTT	1162				
QY	721	CTTCGTCAACACAGGAAACAATGAACTTGTCCAAGAGTTAACAGGCTCTGATTGCTCAGAA	780				
DB	1163	CTTCGTCAACACAGGAAACAATGAACTTGTCCAAGAGTTAACAGGCTCTGATTGCTCAGAA	1222				
QY	781	AGCAATGCAGAGATTGAGAAATTTATCACAAGTTGATGGTCCCTCAAGTGGAGAGCAACTT	840				
DB	1223	AGCAATGCAGAGATTGAGAAATTTATCACAAGTTGATGGTCCCTCAAGTGGAGAGCAACTT	1282				
QY	841	CTTTCAACCCACAGTTTCAAGTGGAGAGGAGGTCTGGGGCATGGAGAAATAACTCA	900				
DB	1283	CTTTCAACCCACAGTTTCAAGTGGAGAGGAGGTCTGGGGCATGGAGAAATAACTCA	1342				
QY	901	TCAGAAATCATCTTTTGCAGATTCTTCTGTAGTTTCAGAAATCAGACACAAAGTTTGGCAGAA	960				
DB	1343	TCAGAAATCATCTTTTGCAGATTCTTCTGTAGTTTCAGAAATCAGACACAAAGTTTGGCAGAA	1402				
QY	961	GGAAGTGTGAGTGTAGATGAAAGTCTTGGACATAACAGCAACATGGGCAGTGATTCA	1020				
DB	1403	GGAAGTGTGAGTGTAGATGAAAGTCTTGGACATAACAGCAACATGGGCAGTGATTCA	1462				

Qy	1021	GGCACCATGGGAAGTGATTTCAGATGAAGAGAAATGTGGCAGCAAGAGCATCCCCGGAGCCA	1080
Db	1463	GGCACCATGGGAAGTGATTTCAGATGAAGAGAAATGTGGCAGCAAGAGCATCCCCGGAGCCA	1522
Qy	1081	GAACTCCAGCTCAGGCCCTTACCAAAATGGAAGTTGCCAGCCAGCCCTTGGAAAGGAAGAAAT	1140
Db	1523	GAACTCCAGCTCAGGCCCTTACCAAAATGGAAGTTGCCAGCCAGCCCTTGGAAAGGAAGAAAT	1582
Qy	1141	ATCATCATCTGCCTCCCTACAGGGAGTGGAAAAACCCAGAGTGGCTGTTTACATTGCCAAG	1200
Db	1583	ATCATCATCTGCCTCCCTACAGGGAGTGGAAAAACCCAGAGTGGCTGTTTACATTGCCAAG	1642
Qy	1201	GATCACCTTAGACAAGAAAGAAAGCATCTGAGCCTGGAAAGTTATAGTTCTTGTCAAT	1260
Db	1643	GATCACCTTAGACAAGAAAGAAAGCATCTGAGCCTGGAAAGTTATAGTTCTTGTCAAT	1702
Qy	1261	AAGGTACTGCTAGTTGAACAGCTCTTCCGCAAGGAGTTCCAACCATTTTTGAAGAAATGG	1320
Db	1703	AAGGTACTGCTAGTTGAACAGCTCTTCCGCAAGGAGTTCCAACCATTTTTGAAGAAATGG	1762
Qy	1321	TATCGTGTATTGGATTAAAGTGGTGATACCCCAACTGAAATATCATTTCCAGAAGTTGTC	1380
Db	1763	TATCGTGTATTGGATTAAAGTGGTGATACCCCAACTGAAATATCATTTCCAGAAGTTGTC	1822
Qy	1381	AAGTCCTGTGATATTATTATCAGTACAGCTCAAAATCCTTGAANAACCTCCCTCTTAAACCTTG	1440
Db	1823	AAGTCCTGTGATATTATTATCAGTACAGCTCAAAATCCTTGAANAACCTCCCTCTTAAACCTTG	1882
Qy	1441	GAAATGGAGAAGATGCTGGTGTTCAAATGTGTCAGACTTTTCCCTCATTTATCATTTGATGAA	1500
Db	1883	GAAATGGAGAAGATGCTGGTGTTCAAATGTGTCAGACTTTTCCCTCATTTATCATTTGATGAA	1942
Qy	1501	TGTCATCACACCAACAAAGAAAGCAGTGTATAATAACATCATGAGGCATTATTTGATGCAG	1560
Db	1943	TGTCATCACACCAACAAAGAAAGCAGTGTATAATAACATCATGAGGCATTATTTGATGCAG	2002
Qy	1561	AAGTTGAAAAACAATAGACTCAAGAAAGAAAAACAACCAGTGATTCCCTTCCCTCAGATA	1620
Db	2003	AAGTTGAAAAACAATAGACTCAAGAAAGAAAAACAACCAGTGATTCCCTTCCCTCAGATA	2062
Qy	1621	CTGGGACTAACAGCTTCACCTGGTGTGGAGGGGCCACGAAGCAAGCCAAAGCTGAAGAA	1680
Db	2063	CTGGGACTAACAGCTTCACCTGGTGTGGAGGGGCCACGAAGCAAGCCAAAGCTGAAGAA	2122
Qy	1681	CACATTTTAAAACTATGTGCCAATCTTGATGCATTTTACTATTAAAACTGTTTAAAGAAAAAC	1740
Db	2123	CACATTTTAAAACTATGTGCCAATCTTGATGCATTTTACTATTAAAACTGTTTAAAGAAAAAC	2182
Qy	1741	CTTGATCAACTGAAAAAACCAATACAGGAGCCCATGCAAGAGTTTGCCATTGCAGATGCA	1800
Db	2183	CTTGATCAACTGAAAAAACCAATACAGGAGCCCATGCAAGAGTTTGCCATTGCAGATGCA	2242
Qy	1801	ACCAGAGAAGATCCATTAAAGAGAAAACTTCTAGAAATTAATGACAAGGATTCAAAACCTTAT	1860
Db	2243	ACCAGAGAAGATCCATTAAAGAGAAAACTTCTAGAAATTAATGACAAGGATTCAAAACCTTAT	2302
Qy	1861	TGTCAAATGAGTCCAATGTTCAGATTTTGGAACTCAACCCCTATGAACAATGGGCCATTCAA	1920
Db	2303	TGTCAAATGAGTCCAATGTTCAGATTTTGGAACTCAACCCCTATGAACAATGGGCCATTCAA	2362
Qy	1921	ATGGAAAAAAAGCTGCAAAAAAGGAAATCGCAAAAGAACGTGTTTGTGCAGAACATTTG	1980
Db	2363	ATGGAAAAAAAGCTGCAAAAAAGGAAATCGCAAAAGAACGTGTTTGTGCAGAACATTTG	2422
Qy	1981	AGGAAGTACAATGAGGCCCTACAAATTAATGACACAATTCGAATGATAGATGCGTATACT	2040
Db	2423	AGGAAGTACAATGAGGCCCTACAAATTAATGACACAATTCGAATGATAGATGCGTATACT	2482
Qy	2041	CATCTTGAAACTTCTATAATGAAGAGAAAAAGATAAGAAAGTTTGCAGTCATAGAAGATGAT	2100
Db	2483	CATCTTGAAACTTCTATAATGAAGAGAAAAAGATAAGAAAGTTTGCAGTCATAGAAGATGAT	2542

QY 2101 AGTGATGAGGGTGGTGATGATGAGTATTGTGATGGTGATGAAGATGAGATGATTTAAAG 2160
Db |||||
2543 AGTGATGAGGGTGGTGATGATGAGTATTGTGATGGTGATGAAGATGAGATGATTTAAAG 2602
QY 2161 AAACCTTTGAAACTGGATGAAACAGATAGATTTCTCATGACTTTTATTTTGGAAACAAT 2220
Db |||||
2603 AAACCTTTGAAACTGGATGAAACAGATAGATTTCTCATGACTTTTATTTTGGAAACAAT 2662
QY 2221 AAAATGTTGAAAGCGTGGCTGAAACCCAGAAATATGAAAAATGAAAGCTGACCAAAATTA 2280
Db |||||
2663 AAAATGTTGAAAGCGTGGCTGAAACCCAGAAATATGAAAAATGAAAGCTGACCAAAATTA 2722
QY 2281 AGAAATACCAATAATGGAGCAATATATACTAGGACTGAGGAATCAGCACGAGGAATAATCTTT 2340
Db |||||
2723 AGAAATACCAATAATGGAGCAATATATACTAGGACTGAGGAATCAGCACGAGGAATAATCTTT 2782
QY 2341 ACAAAACACGACAGAGTGCAATATGGGCTTTCCAGTGGAATTACTGAAATGAAAAATTT 2400
Db |||||
2783 ACAAAACACGACAGAGTGCAATATGGGCTTTCCAGTGGAATTACTGAAATGAAAAATTT 2842
QY 2401 GCTGAAGTAGGAGTCAAAAGCCCAACCATCTGATTGGAGCTGGACACAGCAGTGAGTTCAAA 2460
Db |||||
2843 GCTGAAGTAGGAGTCAAAAGCCCAACCATCTGATTGGAGCTGGACACAGCAGTGAGTTCAAA 2902
QY 2461 CCCATGACACAGAAATGAACAAAGAAAGTCATTAGTAAATTTGCGCACTGGAAAAATCAAT 2520
Db |||||
2903 CCCATGACACAGAAATGAACAAAGAAAGTCATTAGTAAATTTGCGCACTGGAAAAATCAAT 2962
QY 2521 CTGCTTATCGTACACAGTGGCAGAAAGAGGTCTGGATATTAAGAATGTAACATGTT 2580
Db |||||
2963 CTGCTTATCGTACACAGTGGCAGAAAGAGGTCTGGATATTAAGAATGTAACATGTT 3022
QY 2581 ATCCGTTATGGTCTCGTCACCAATGAAATAGCCATGGTCCAGSCCGTGGTCGAGCCAGA 2640
Db |||||
3023 ATCCGTTATGGTCTCGTCACCAATGAAATAGCCATGGTCCAGSCCGTGGTCGAGCCAGA 3082
QY 2641 GCTGATGAGAGCACCTACGTCTGTTGCTCACAGTGGTTCAGGAGTTATCGAACATGAG 2700
Db |||||
3083 GCTGATGAGAGCACCTACGTCTGTTGCTCACAGTGGTTCAGGAGTTATCGAACATGAG 3142
QY 2701 ACAGTTAATGATTTCCGAGAGAAAGATGATGTATAAAGCTATACATTTGTTCAAAATATG 2760
Db |||||
3143 ACAGTTAATGATTTCCGAGAGAAAGATGATGTATAAAGCTATACATTTGTTCAAAATATG 3202
QY 2761 AAACAGAGGAGTATGCTCATAAAGATTTTGGAAATTACAGATGCAAAAGTATAATGGAAAG 2820
Db |||||
3203 AAACAGAGGAGTATGCTCATAAAGATTTTGGAAATTACAGATGCAAAAGTATAATGGAAAG 3262
QY 2821 AAAATGAAACCAAGAGAAATATTGCCAAAGCATTAAGAATAACCCATCACTAATAACT 2880
Db |||||
3263 AAAATGAAACCAAGAGAAATATTGCCAAAGCATTAAGAATAACCCATCACTAATAACT 3322
QY 2881 TTCCCTTTGCAAAAACTGCAGTGTGCTAGCCTGTTCTGGGGAAGATATCCATGTAATTGAG 2940
Db |||||
3323 TTCCCTTTGCAAAAACTGCAGTGTGCTAGCCTGTTCTGGGGAAGATATCCATGTAATTGAG 3382
QY 2941 AAAATGCATCACGTCAATATGACCCCAAGAAATTCAGGAACTTTACATTTGTAAGAGAAAC 3000
Db |||||
3383 AAAATGCATCACGTCAATATGACCCCAAGAAATTCAGGAACTTTACATTTGTAAGAGAAAC 3442
QY 3001 AAAGCACTGGGAAACAATGATGGTGCACAAAGGCTTAGATTTGCTTGTCTCAAAATA 3060
Db |||||
3443 AAACACTGCAAAAGAAGTGTGCCGACTATCAATAAATGGTGAATCATCTGCAAAATGT 3502
QY 3061 GGCCAGGCTTTGGGAAACAATGATGGTGCACAAAGGCTTAGATTTGCTTGTCTCAAAATA 3120
Db |||||
3503 GGCCAGGCTTTGGGAAACAATGATGGTGCACAAAGGCTTAGATTTGCTTGTCTCAAAATA 3562
QY 3121 AGGAATTTTGTAGTGGTTTTCAAAAATAATTTCAACAAAGAAACAATACAAAAAGTGGGTA 3180
Db |||||
3563 AGGAATTTTGTAGTGGTTTTCAAAAATAATTTCAACAAAGAAACAATACAAAAAGTGGGTA 3622
QY 3181 GAATTACCTATCACATTTCCCAATCTTGACTATTTCAGAAATGCTGTTATTTTAGTGATGAG 3240

Db 3623 GAATTACCTATCACATTTCCCAATCTTGACTATTCAGAAATGCTGTTATTTAGTGATGAG 3682
QY 3241 GATTAGCACTTGATTGAAGATTCTTTTAAATAACTATCAGTTAAACAATTAATATGATTA 3300
Db |||||
3683 GATTAGCACTTGATTGAAGATTCTTTTAAATAACTATCAGTTAAACAATTAATATGATTA 3742
QY 3301 TGATTAAATGATTATTCATTATGCTACAGAACTGACATAAAGAATCAATAAAATGATTGTTTTA 3360
Db |||||
3743 TGATTAAATGATTATTCATTATGCTACAGAACTGACATAAAGAATCAATAAAATGATTGTTTTA 3802
QY 3361 CTCTG 3365
Db |||||
3803 CTCTG 3807

RESULT 6
US-10-055-475-4
; Sequence 4, Application US/10055475
; Publication No. US20030022855A1
; GENERAL INFORMATION:
; APPLICANT: Fisher, Paul B.
; APPLICANT: Kang, Dong-Chul
; APPLICANT: Gopalkrishnan, Rahul V.
; TITLE OF INVENTION: USE OF MDA-5 AS AN ANTIVIRAL AND
; TITLE OF INVENTION: ANTIPROLIFERATIVE AGENT
; FILE REFERENCE: A34614-A-PCT-USA-A (070050.1921)
; CURRENT APPLICATION NUMBER: US/10/055,475
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: PCT/US01/06960
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 09/515,363
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 3627
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-055-475-4

Query Match 99.8%; Score 3358.6; DB 14; Length 3627;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3361; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GCGCGCCGGCCTGAGAGCCCTGTGGACAACCTCGTCATTGTTCAGGCACAGAGCGGTAGAC 60
Db |||||
1 GCGCGCCGGCCTGAGAGCCCTGTGGACAACCTCGTCATTGTTCAGGCACAGAGCGGTAGAC 60
QY 61 CCTGCTTCTCTAAGTGGCAGCGGACAGCGGCACGACATTTACCTGTCCCAGACAA 120
Db |||||
61 CCTGCTTCTCTAAGTGGCAGCGGACAGCGGCACGACATTTACCTGTCCCAGACAA 120
QY 121 CAGCACCATCTGCTTGGGAGAACCTCTCCCTTCTCTGAGAAAGAAAGATGCGAATGGG 180
Db |||||
121 CAGCACCATCTGCTTGGGAGAACCTCTCCCTTCTCTGAGAAAGAAAGATGCGAATGGG 180
QY 181 TATTCACAGACGAGAAATTCGCGCTATCTCATCTCGTGTTCAGGGCCAGGGTGAAATG 240
Db |||||
181 TATTCACAGACGAGAAATTTCCGCTATCTCATCTCGTGTTCAGGGCCAGGGTGAAATG 240
QY 241 TACATCCAGGTGGAGCCCTGTGCTGGACTACCTGACCTTTCTGCTGCAGAGGTGAAGGAG 300
Db |||||
241 TACATCCAGGTGGAGCCCTGTGCTGGACTACCTGACCTTTCTGCTGCAGAGGTGAAGGAG 300
QY 301 CAGATTCAGAGGACAGTCGCCACCTCCGGAACTGCGAGGCAGTTGAACTGCTGAGC 360
Db |||||
301 CAGATTCAGAGGACAGTCGCCACCTCCGGAACTGCGAGGCAGTTGAACTGCTGAGC 360
QY 361 ACCTTGGAGAGGGAGTCTGGACCTTGGTTGGACTCGGGAATTCGTGGAGGCCCTCCGG 420
Db |||||
361 ACCTTGGAGAGGGAGTCTGGACCTTGGTTGGACTCGGGAATTCGTGGAGGCCCTCCGG 420

QY 421 AGAACGGCAGCCCTCTGGCCGCGCCGCTACATGAACCCCTGAGCTACGGACTTGCCCTCT 480
Db |||||
QY 421 AGAACGGCAGCCCTCTGGCCGCGCCGCTACATGAACCCCTGAGCTACGGACTTGCCCTCT 480
Db |||||
QY 481 CCATCGTTTGAGAACGCTCATGATGAATATCTCCAACTGCTGAACCTCTTTCAGCCCACT 540
Db |||||
QY 541 CTGGTGGACAAAGCTTCTAGTTAGAGACGCTCTGGATAAGTGCATGGAGGAGAACTGTTG 600
Db |||||
QY 601 ACAATTGAAGACAGAAACCGGATTGCTGCTGCAGAAACAATGGAATGAATCAGGTGTA 660
Db |||||
QY 661 AGAGAGCTACTAAAAGGATTGTGCAGAAAGAAACTGGTTCCTGCAATTTCTGAATGTT 720
Db |||||
QY 721 CTTTCTCAACACAGGAAACAATGAACCTTGTCGAAGATTAAACAGGCTCTGATTCAGAA 780
Db |||||
QY 781 AGCAATGCAGAGATTGAGAAATTTATCACAAGTTGATGGTCTCAAGTGGAAAGCAACTT 840
Db |||||
QY 841 CTTTCAACCCACAGTTTCAGCCAAATCTGGAGAAGGAGTCTGGGSCATGGAGATAACTCA 900
Db |||||
QY 901 TCAGAAATCATCTTTTGAGATTCTTCTGTAGTTTCAGAAATCAGACACAAGTTTGGCAGAA 960
Db |||||
QY 961 GGAAGTGTGAGCTGCTTAGATGAAAGTCTTGGACATAACAGCAACATGGGCAGTGATTCA 1020
Db |||||
QY 1021 GGCACCATGGGAAGTGATTCAGATGAAGAAATGTGGCAGCAAGAGCATCCCCGGAGCCA 1080
Db |||||
QY 1081 GAACTCCAGCTCAGGCCTTACCAAATGGAAGTTGCCAGCCAGCTTGGAAAGGAAGAT 1140
Db |||||
QY 1141 ATCATCATCTGCCTCCCTACAGGGAGTGGAAACCCAGAGTGGTGTTCATTTGCCAAG 1200
Db |||||
QY 1201 GATCATTAGACAAGAAAGAAAAGCATCTGAGCCTGGAAAGTTATAGTTCTTGTCAAT 1260
Db |||||
QY 1261 AAGGTACTGCTAGTTGAACAGCTCTTCGCAAGGAGTTCCAACCATTTTGAAGAAATGG 1320
Db |||||
QY 1321 TATCGTGTATTGGATTAAAGTGGTGATACCCAACTGAAATATCATTTCCAGAAGTTGTC 1380
Db |||||
QY 1381 AAGTCCCTGTGATATTATTATCAGTACAGCTCAAACTCCTTGAAACTCCCTCTTAAACTTG 1440
Db |||||
QY 1441 GAAATGGAGAGATGCTGGTTCATTTGTTCAGACTTTTCCCTCATTTATCATTTGATGAA 1500
Db |||||
QY 1501 TGTTCATCACACCAACAAAGAACGAGTGATATAATAACATCATGAGGCATTATTTGATGCAG 1560

Db |||||
QY 1501 TGTTCATCACCAACAAAGAACGAGTGTATAATAACATCATGAGGCATTATTTGATGCAG 1560
Db |||||
QY 1561 AAGTTGAAAAACAATAGACTCAAGAAAGAAAAACAACCCAGTGATTTCCCTTCAGATA 1620
Db |||||
QY 1621 CTGGGACTAAACAGCTTTCACCTGGTGTGGAGGGGCCACGAAGCAAGCCAAAGCTGAAGAA 1680
Db |||||
QY 1681 CACATTTTAAAACTATGTGCCAATCTTGATGCATTTTACTATTAAAACTGTTAAAGAAAC 1740
Db |||||
QY 1741 CTTGATCAACTGAAAAACCAAAATACAGGAGCCATGCAAGAGTTTGCCATTGCAGATGCA 1800
Db |||||
QY 1801 ACCAGAGAAGATCCATTTTAAAGAGAAACTTCTAGAAATAATGACAAGGATTCAAAACCTAT 1860
Db |||||
QY 1861 TGTCAAATGAGTCCAATGTTCAGATTTTGGAACTCAACCCCTATGAACAATGGGCCATTCAA 1920
Db |||||
QY 1921 ATGAAAAAAAGCTGCAAAAAAAGGAAATCGAAAAAGACGTGTTTGTGCAGAACATTTG 1980
Db |||||
QY 1981 AGGAAGTACAATGAGGCCCTACAAATTAATGACACAAATTCGAATGATAGATCGGTACT 2040
Db |||||
QY 2041 CATCTTGAAACTTCTATAATGAAGAGAAAGATAAGAAAGTTTGCAGTCATAGAAGATGAT 2100
Db |||||
QY 2101 AGTGATGAGGGTGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2160
Db |||||
QY 2161 AAACTTTGAAAAGCTGGATGAAAACAGATAGATTTCTCATGACTTTATTTTGTGAAAAACAAT 2220
Db |||||
QY 2221 AAAATGTTGAAAAAGCTGGCTGAAAAACCCAGAAATATGAAAAATGAAAAAGCTGACCAAAATTA 2280
Db |||||
QY 2281 AGAAATACCAATAATGGAGCAATATACTAGGACTGAGGAATCAGCACGAGGAATAATCTTT 2340
Db |||||
QY 2341 ACAAAAACACGACAGAGTGCAATATGCGCTTTCCCAGTGGATTACTGAAAAATGAAAAATTT 2400
Db |||||
QY 2401 GCTGAAGTAGGAGTCAAAGCCCAACATCTGATTGGAGCTGGAACACAGCAGTGAAGTCAA 2460
Db |||||
QY 2461 CCCATGACACAGAAATGAACAAAAAGAAAGTCAATAGTAAATTTCCACCTGGAATAATCAAT 2520
Db |||||
QY 2521 CTGCTTATCGCTACCAAGTGGCAGAAAGGCTCTGGATATTAAAGAAATGTAACATTTGTT 2580
Db |||||
QY 2581 ATCCGTTATGGTCTCGTCACCAATGAATAGCCATGGTCCAGGCCCGTGGTCGAGCCAGA 2640

Db 2581 ATCCGTTATGTTCTCGTCACCAATGAATAGCCATGGTCCAGGCCCGTGGTCGAGCCAGA 2640

Qy 2641 GCTGATGAGACCACTACGTCCTGGTTGCTCAGAGTGGTTTCAGGAGTTATCGAACATGAG 2700

Db 2641 GCTGATGAGACCACTACGTCCTGGTTGCTCAGAGTGGTTTCAGGAGTTATCGAACGTGAG 2700

Qy 2701 ACAGTTAATGATTTCCGAGAGAGATGATGATAAAGCTATACATTTGTTCAAAATATG 2760

Db 2701 ACAGTTAATGATTTCCGAGAGAGATGATGATAAAGCTATACATTTGTTCAAAATATG 2760

Qy 2761 AAACCCAGAGGATGCTCATAGATTTTGGAAATTACAGATGCAAAAGTATATGGAAG 2820

Db 2761 AAACCCAGAGGATGCTCATAGATTTTGGAAATTACAGATGCAAAAGTATATGGAAG 2820

Qy 2821 AAAATGAACCAAGAGAAATATTGCCAAGCATTACAAGATAACCCATCACAATAACT 2880

Db 2821 AAAATGAACCAAGAGAAATATTGCCAAGCATTACAAGATAACCCATCACAATAACT 2880

Qy 2881 TTCCCTTTGCAAAAACTGCAGTGTGCTAGCCTGTTCTGGGGAAGATATCCATGTAATTGAG 2940

Db 2881 TTCCCTTTGCAAAAACTGCAGTGTGCTAGCCTGTTCTGGGGAAGATATCCATGTAATTGAG 2940

Qy 2941 AAAATGCATCAGTCAATATGACCCCAAGATTTCAAGGAATTTACATTTGTAAGAGAAAAC 3000

Db 2941 AAAATGCATCAGTCAATATGACCCCAAGATTTCAAGGAATTTACATTTGTAAGAGAAAAC 3000

Qy 3001 AAAGCACTGCAAAAGAAAGTGTGCCGACTATCAAAATAAATGGTGAATCATCTGCAAAATGT 3060

Db 3001 AAAGCACTGCAAAAGAAAGTGTGCCGACTATCAAAATAAATGGTGAATCATCTGCAAAATGT 3060

Qy 3061 GGCCAGGCTTGGGGAACAATGATGGTGCAAAAGGCTTAGATTTGCCTTGTCTCAAAATA 3120

Db 3061 GGCCAGGCTTGGGGAACAATGATGGTGCAAAAGGCTTAGATTTGCCTTGTCTCAAAATA 3120

Qy 3121 AGGAATTTTGTAGTGGTTTTTCAAAATAAATTCAACAAAGAAACAATACAAAAAGTGGTA 3180

Db 3121 AGGAATTTTGTAGTGGTTTTTCAAAATAAATTCAACAAAGAAACAATACAAAAAGTGGTA 3180

Qy 3181 GAATTACCTATCAGATTTCCCAATCTTGACTATTTCAGAAATGCTGTTTATTAGTGATGAG 3240

Db 3181 GAATTACCTATCAGATTTCCCAATCTTGACTATTTCAGAAATGCTGTTTATTAGTGATGAG 3240

Qy 3241 GATTAGCACTTGATGAAGATTTCTTTTAAATACTATCAGTTAAACATTTAAATATGATTA 3300

Db 3241 GATTAGCACTTGATGAAGATTTCTTTTAAATACTATCAGTTAAACATTTAAATATGATTA 3300

Qy 3301 TGATTAATGATTCATATGCTACAGAACTGACATAAGAAATCAATAAAATGATTGTTT 3360

Db 3301 TGATTAATGATTCATATGCTACAGAACTGACATAAGAAATCAATAAAATGATTGTTT 3360

Qy 3361 CTCTG 3365

Db 3361 CTCTG 3365

RESULT 7

US-10-723-860-5057

; Sequence 5057, Application US/10723860

; Publication No. US20040253606A1

; GENERAL INFORMATION:

; APPLICANT: Aziz, Natasha

; APPLICANT: Ginsburg, Wendy M.

; APPLICANT: Zlotnik, Albert

; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &

; TITLE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators

; FILE REFERENCE: 05882.0193.NPUS01

; CURRENT APPLICATION NUMBER: US/10/723,860

; CURRENT FILING DATE: 2003-11-26

; PRIOR APPLICATION NUMBER: 60/429,739

; PRIOR FILING DATE: 2002-11-26

; NUMBER OF SEQ ID NOS: 8393

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 5057

; LENGTH: 3668

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (719)..(739)

; OTHER INFORMATION: n is a, c, g, or t

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (3426)..(3431)

; OTHER INFORMATION: n is a, c, g, or t

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (3434)..(3440)

; OTHER INFORMATION: n is a, c, g, or t

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (3442)..(3442)

; OTHER INFORMATION: n is a, c, g, or t

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (3444)..(3444)

; OTHER INFORMATION: n is a, c, g, or t

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (3448)..(3448)

; OTHER INFORMATION: n is a, c, g, or t

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (3455)..(3455)

; OTHER INFORMATION: n is a, c, g, or t

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (3462)..(3462)

; OTHER INFORMATION: n is a, c, g, or t

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (3473)..(3473)

; OTHER INFORMATION: n is a, c, g, or t

US-10-723-860-5057

Query Match 99.2%; Score 3339.2; DB 18; Length 3668;

Best Local Similarity 99.3%; Pred. No. 0;

Matches 3341; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

Qy 1 GCGCGCGGCTGAGAGCCCTGTGGACAACCTCGTCATTGTTCAGGCACAGAGCGGTAGAC 60

Db 1 GCGCGCGGCTGAGAGCCCTGTGGACAACCTCGTCATTGTTCAGGCACAGAGCGGTAGAC 60

Qy 61 CCTGCTTCTTAAGTGGCAGCGGACAGCGGCACGACATTTACCTGTCCCGCAGACAA 120

Db 61 CCTGCTTCTTAAGTGGCAGCGGACAGCGGCACGACATTTACCTGTCCCGCAGACAA 120

Qy 121 CAGCACCATCTGCTTGGGAAACCCCTCTCCCTTCTCTGAGAAAGATGTCGAATGGG 180

Db 121 CAGCACCATCTGCTTGGGAAACCCCTCTCCCTTCTCTGAGAAAGATGTCGAATGGG 180

Qy 181 TATTCCACAGACGAGAAATTTCCGCTATCTCATCTCGTCTTCAGGGCCAGGTTGAAAATG 240

Db 181 TATTCCACAGACGAGAAATTTCCGCTATCTCATCTCGTCTTCAGGGCCAGGTTGAAAATG 240

Qy 241 TACATCCAGGTGGAGCCCTGTGCTGGACTACCTGACCTTTCTGCCTGCAGAGGTGAAGGAG 300

Db 241 TACATCCAGGTGGAGCCCTGTGCTGGACTACCTGACCTTTCTGCCTGCAGAGGTGAAGGAG 300

Qy 301 CAGATTCAGAGGACAGTCGCCACCTCCGGGAACATGCAGGCAGTTGAACTGCTGAGC 360

Db 301 CAGATTCAGAGGACAGTCGCCACCTCCGGGAACATGCAGGCAGTTGAACTGCTGAGC 360

Qy 361 ACCTTGGAGAAAGGAGTCTGSCACCTTGGTTGGACTCGGGAATTCGTGGAGCCCTCCGG 420

Db 361 ACCTTGGAGAAAGGAGTCTGSCACCTTGGTTGGACTCGGGAATTCGTGGAGCCCTCCGG 420

QY 421 AGAACGGCAGCCCTCTGGCCGCCGCTACATGAACCCCTGAGCTCACGGACTTGCCCTCT 480
Db |||||
QY 421 AGAACGGCAGCCCTCTGGCCGCCGCTACATGAACCCCTGAGCTCACGGACTTGCCCTCT 480
Db |||||
QY 481 CCATCGTTTGAGAACGCTCATGATGAATATCTCAACTGCTGAACCTCTTCAGCCCACT 540
Db |||||
QY 481 CCATCGTTTGAGAACGCTCATGATGAATATCTCAACTGCTGAACCTCTTCAGCCCACT 540
Db |||||
QY 541 CTGGTGGACAAGCTTCTAGTTAGAGACGCTCTGGATAAGTGCTATGGAGGAGAACTGTTG 600
Db |||||
QY 541 CTGGTGGACAAGCTTCTAGTTAGAGACGCTCTGGATAAGTGCTATGGAGGAGAACTGTTG 600
Db |||||
QY 601 ACAATTGAAGACAGAAACCGGATTGCTGCTGCAGAAACAATGGAATGAATCAGGTGTA 660
Db |||||
QY 601 ACAATTGAAGACAGAAACCGGATTGCTGCTGCAGAAACAATGGAATGAATCAGGTGTA 660
Db |||||
QY 661 AGAGAGCTACTAAAAAGGATTGTGCAGAAAGAAACTGGTTCTCTGCATTTCTGAATGTT 720
Db |||||
QY 661 AGAGAGCTACTAAAAAGGATTGTGCAGAAAGAAACTGGTTCTCTGCATTTCTGAATGNN 720
Db |||||
QY 721 CTTCTGTCAAAACAGGAAACAATGAACCTTGTCGAAGAGTTAACAGGCTCTGATGCTCAGAA 780
Db |||||
QY 721 NNNNNNNNNNNNNNNNATGAACCTTGTCGAAGAGTTAACAGGCTCTGATGCTCAGAA 780
Db |||||
QY 781 AGCAATGCAGAGATTGAGAAATTTATCACAAGTTGATGGTCTCAAGTGGAGGCAACTT 840
Db |||||
QY 781 AGCAATGCAGAGATTGAGAAATTTATCACAAGTTGATGGTCTCAAGTGGAGGCAACTT 840
Db |||||
QY 841 CTTTCAACCACAGTTTCAGCCAAATCTGGAGAGGAGGTCTGGGCGATGGAGATAACTCA 900
Db |||||
QY 841 CTTTCAACCACAGTTTCAGCCAAATCTGGAGAGGAGGTCTGGGCGATGGAGATAACTCA 900
Db |||||
QY 901 TCAGAAATCATCTTTTCAGAGATTCTTCTGTAGTTTCAGAAATCAGACACAAGTTTGGCAGAA 960
Db |||||
QY 901 TCAGAAATCATCTTTTCAGAGATTCTTCTGTAGTTTCAGAAATCAGACACAAGTTTGGCAGAA 960
Db |||||
QY 961 GGAAGTGTGAGTCTGATGAAAGTCTTGGAATAACAGCAATGAGCAATGAGCAATGAGTCA 1020
Db |||||
QY 961 GGAAGTGTGAGTCTGATGAAAGTCTTGGAATAACAGCAATGAGCAATGAGCAATGAGTCA 1020
Db |||||
QY 1021 GGCACCATGGGAAGTGAATTCAGATGAAGAAATGTGGCAGCAAGAGCATCCCCGGAGCCA 1080
Db |||||
QY 1021 GGCACCATGGGAAGTGAATTCAGATGAAGAAATGTGGCAGCAAGAGCATCCCCGGAGCCA 1080
Db |||||
QY 1081 GAACTCCAGCTCAGGCCTTACCAAATGGAAGTTGCCAGCCAGCCTTGAAGGGAAGAAAT 1140
Db |||||
QY 1081 GAACTCCAGCTCAGGCCTTACCAAATGGAAGTTGCCAGCCAGCCTTGAAGGGAAGAAAT 1140
Db |||||
QY 1141 ATCATCATCTGCCTCCCTACAGGGAGTGGAACCAACAGAGTGGTGTTCATTTGCCAAG 1200
Db |||||
QY 1141 ATCATCATCTGCCTCCCTACAGGGAGTGGAACCAACAGAGTGGTGTTCATTTGCCAAG 1200
Db |||||
QY 1201 GATCACTTAGACAAGAAAGAAAGCATCTGAGCCTGGAAAGTTATAGTTCTTGTCAAT 1260
Db |||||
QY 1201 GATCACTTAGACAAGAAAGAAAGCATCTGAGCCTGGAAAGTTATAGTTCTTGTCAAT 1260
Db |||||
QY 1261 AAGGTACTGCTAGTTGAAACAGCTCTTCCGCAAGGAGTTCCAAACCATTTTGAAGAAATGG 1320
Db |||||
QY 1261 AAGGTACTGCTAGTTGAAACAGCTCTTCCGCAAGGAGTTCCAAACCATTTTGAAGAAATGG 1320
Db |||||
QY 1321 TATCGTGTATTGGATTAAAGTGGTATACCCAACTGAAATATCATTTCCAGAAGTTGTC 1380
Db |||||
QY 1321 TATCGTGTATTGGATTAAAGTGGTATACCCAACTGAAATATCATTTCCAGAAGTTGTC 1380
Db |||||
QY 1381 AAGTCCCTGTGATATTATTATCAGTACAGCTCAAAATCCTTGAAACTCCCTCTTAAACTTG 1440
Db |||||
QY 1381 AAGTCCCTGTGATATTATTATCAGTACAGCTCAAAATCCTTGAAACTCCCTCTTAAACTTG 1440
Db |||||
QY 1441 GAAAAATGGAGAGATGCTGGTGTCAATTGTGAGACTTTTCCCTCATTTATCATTTGATGAA 1500
Db |||||
QY 1441 GAAAAATGGAGAGATGCTGGTGTCAATTGTGAGACTTTTCCCTCATTTATCATTTGATGAA 1500
Db |||||
QY 1501 TGTTCATCACACCAAGAAAGCAGTGATATAATAACATCATGAGGCATTATTTGATGCAG 1560
Db |||||

Db |||||
QY 1501 TGTTCATCACACCAACAAAGAAAGCAGTGTATAATAACATCATGAGGCATTATTTGATGCAG 1560
Db |||||
QY 1561 AAGTTGAAAAACAATAGACTCAAGAAAGAAAAACAACCAAGTATTCCTTCAGATA 1620
Db |||||
QY 1561 AAGTTGAAAAACAATAGACTCAAGAAAGAAAAACAACCAAGTATTCCTTCAGATA 1620
Db |||||
QY 1621 CTGGGACTAACAGCTTTCACCTGGTGTGGAGGGCCACGAAGCAAGCTGAAGAA 1680
Db |||||
QY 1621 CTGGGACTAACAGCTTTCACCTGGTGTGGAGGGCCACGAAGCAAGCTGAAGAA 1680
Db |||||
QY 1681 CACATTTTAAAACTATGTGCCAATCTTGATGCAATTTACTATTAAAACTGTTAAAGAAAAAC 1740
Db |||||
QY 1681 CACATTTTAAAACTATGTGCCAATCTTGATGCAATTTACTATTAAAACTGTTAAAGAAAAAC 1740
Db |||||
QY 1741 CTTGATCAACTGAAAAACCAAAATACAGGAGCCATGCAAGAACTTGCCATTGCAGATGCA 1800
Db |||||
QY 1741 CTTGATCAACTGAAAAACCAAAATACAGGAGCCATGCAAGAACTTGCCATTGCAGATGCA 1800
Db |||||
QY 1801 ACCAGAGAAGATCCATTTTAAAGAGAAACTTCTAGAAATAATGACAAGGATTCAAAATTAT 1860
Db |||||
QY 1801 ACCAGAGAAGATCCATTTTAAAGAGAAACTTCTAGAAATAATGACAAGGATTCAAAATTAT 1860
Db |||||
QY 1861 TGTCAAATGAGTCCAAATGTGAGATTTTGGAACTCAAACCTATGAACAAATGGGCCATTCAA 1920
Db |||||
QY 1861 TGTCAAATGAGTCCAAATGTGAGATTTTGGAACTCAAACCTATGAACAAATGGGCCATTCAA 1920
Db |||||
QY 1921 ATGAAAAAAAAGCTGCAAAAAAGGAAATCGCAAGAACGTGTTTGTGCAGAACATTG 1980
Db |||||
QY 1921 ATGAAAAAAAAGCTGCAAAAAAGGAAATCGCAAGAACGTGTTTGTGCAGAACATTG 1980
Db |||||
QY 1981 AGGAAGTACAATGAGGCCCTACAAATTAATGACACAATTCGAATGATAGATGCGTATACT 2040
Db |||||
QY 1981 AGGAAGTACAATGAGGCCCTACAAATTAATGACACAATTCGAATGATAGATGCGTATACT 2040
Db |||||
QY 2041 CATCTTGAAACTTTCTATAATGAAGAGAAAGATAAGAACTTGCAGTCATAGAAGATGAT 2100
Db |||||
QY 2041 CATCTTGAAACTTTCTATAATGAAGAGAAAGATAAGAACTTGCAGTCATAGAAGATGAT 2100
Db |||||
QY 2101 AGTGATGAGGGTGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2160
Db |||||
QY 2101 AGTGATGAGGGTGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2160
Db |||||
QY 2161 AAACCTTTGAAACTGGATGAAACAGATAGATTTCTCATGACTTATTTTGTGAAAAACAAT 2220
Db |||||
QY 2161 AAACCTTTGAAACTGGATGAAACAGATAGATTTCTCATGACTTATTTTGTGAAAAACAAT 2220
Db |||||
QY 2221 AAAATGTTGAAAAAGCTGGCTGAAAAACCCAGAAATATGAAAAATGAAAGCTGACCAAAATTA 2280
Db |||||
QY 2221 AAAATGTTGAAAAAGCTGGCTGAAAAACCCAGAAATATGAAAAATGAAAGCTGACCAAAATTA 2280
Db |||||
QY 2281 AGAAATACCAATAATGGAGCAATATACTAGGACTGAGGAATCAGCACGAGGAATAATCTTT 2340
Db |||||
QY 2281 AGAAATACCAATAATGGAGCAATATACTAGGACTGAGGAATCAGCACGAGGAATAATCTTT 2340
Db |||||
QY 2341 ACAAACACGACAGAGTGCAATATGCGCTTTCCCAGTGGATTTACTGAAAAATGAAAAATTT 2400
Db |||||
QY 2341 ACAAACACGACAGAGTGCAATATGCGCTTTCCCAGTGGATTTACTGAAAAATGAAAAATTT 2400
Db |||||
QY 2401 GCTGAAGTAGGAGTCAAAGCCCACTCTGATTTGGAGCTGGACACAGCAGTGAGTTCAA 2460
Db |||||
QY 2401 GCTGAAGTAGGAGTCAAAGCCCACTCTGATTTGGAGCTGGACACAGCAGTGAGTTCAA 2460
Db |||||
QY 2461 CCCATGACACAGAAATGAACAAAAAGAAAGTCAATAGTAAATTTGCACTGGAAAAATCAAT 2520
Db |||||
QY 2461 CCCATGACACAGAAATGAACAAAAAGAAAGTCAATAGTAAATTTGCACTGGAAAAATCAAT 2520
Db |||||
QY 2521 CTGCTTATCGCTACCAAGTGGCAGAAAGGCTCTGGATATTAAAGAAATGTAACATTGTT 2580
Db |||||
QY 2521 CTGCTTATCGCTACCAAGTGGCAGAAAGGCTCTGGATATTAAAGAAATGTAACATTGTT 2580
Db |||||
QY 2581 ATCCGTTATGCTCGTCACCAATGAATAGCCATGGTCCAGGCCGCTGGTCGAGCCAGA 2640
Db |||||

```
Db 2581 ATCCGTTATGGTCTCGTCACCAATGAANAATAGCCATGGTCCAGGCCCGTGGTCGAGCCAGA 2640
QY 2641 GCTGATGAGAGCACCTACGTCCTGGTTGCTCAGAGTGGTTCAGGAGTTATCGAACATGAG 2700
Db 2641 GCTGATGAGAGCACCTACGTCCTGGTTGCTCAGAGTGGTTCAGGAGTTATCGAACGTTGAG 2700
QY 2701 ACAGTTAATGATTTCCGAGAGAGAGATGATGTATAAAGCTATACATTTGTGTTCAAAAATATG 2760
Db 2701 ACAGTTAATGATTTCCGAGAGAGAGATGATGTATAAAGCTATACATTTGTGTTCAAAAATATG 2760
QY 2761 AAACCAGAGGAGTATGCTCATAGATTTTGGAAATTACAGATGCAAGATATAATGGAAGAAG 2820
Db 2761 AAACCAGAGGAGTATGCTCATAGATTTTGGAAATTACAGATGCAAGATATAATGGAAGAAG 2820
QY 2821 AAAATGAAACCAAGAGAAATATTGCCAAGCATTACAAGAAATAACCCATCACTAATAACT 2880
Db 2821 AAAATGAAACCAAGAGAAATATTGCCAAGCATTACAAGAAATAACCCATCACTAATAACT 2880
QY 2881 TTCCTTTGCAAAAACCTGCAGTGTGCTAGCCTGTTCTGGGGAAGATATCCATGTAATTGAG 2940
Db 2881 TTCCTTTGCAAAAACCTGCAGTGTGCTAGCCTGTTCTGGGGAAGATATCCATGTAATTGAG 2940
QY 2941 AAAATGCATCAGTCAATATGACCCCAAGATTTCAAGGAATTTACATTTGTAAGAGAAAAC 3000
Db 2941 AAAATGCATCAGTCAATATGACCCCAAGATTTCAAGGAATTTACATTTGTAAGAGAAAAC 3000
QY 3001 AAAGCACTGCAAAAGAAAGTGTGCCGACTATCAAAATAAATGGTGAAATCATCTGCAAAATGT 3060
Db 3001 AAAGCACTGCAAAAGAAAGTGTGCCGACTATCAAAATAAATGGTGAAATCATCTGCAAAATGT 3060
QY 3061 GGCCAGGCTTGGGGAACAATGATGGTGCAAAAAGGCTTAGATTTGCCCTTGCTCAAAAATA 3120
Db 3061 GGCCAGGCTTGGGGAACAATGATGGTGCAAAAAGGCTTAGATTTGCCCTTGCTCAAAAATA 3120
QY 3121 AGGAATTTTGTAGTGGTTTCAAAAATAAATTCAACAAGAAACAATAACAAGAGTGGGTA 3180
Db 3121 AGGAATTTTGTAGTGGTTTCAAAAATAAATTCAACAAGAAACAATAACAAGAGTGGGTA 3180
QY 3181 GAATTACCTATCACATTTCCCAATCTTGACTATTTCAGAAATGCTGTTATTATTAGTATGAG 3240
Db 3181 GAATTACCTATCACATTTCCCAATCTTGACTATTTCAGAAATGCTGTTATTATTAGTATGAG 3240
QY 3241 GATTAGCACTTGATTGAAGATTCCTTTAAATACTACTATCAGTTAAACATTTAATATGATTA 3300
Db 3241 GATTAGCACTTGATTGAAGATTCCTTTAAATACTACTATCAGTTAAACATTTAATATGATTA 3300
QY 3301 TGATTAATGATTTCAATATGCTACAGAACTGACATAAGAAATCAATAAAATGATTGTTTTA 3360
Db 3301 TGATTAATGATTTCAATATGCTACAGAACTGACATAAGAAATCAATAAAATGATTGTTTTA 3360
QY 3361 CTCTG 3365
Db 3361 CTCTG 3365

RESULT 8
US-10-275-822A-1
; Sequence 1, Application US/10275822A
; Publication No. US20040086500A1
; GENERAL INFORMATION:
; APPLICANT: Bahr, Georges
; APPLICANT: Cocude, Cecile
; APPLICANT: Capron, Andre
; TITLE OF INVENTION: RH116 Polypeptide and its Fragments and Polynucleotides Encoding
; TITLE OF INVENTION: said Polypeptides and Therapeutic Uses
; FILE REFERENCE: 017753-170
; CURRENT APPLICATION NUMBER: US/10/275,822A
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: PCT/FR01/01441
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: FR 00/06,030
; PRIOR FILING DATE: 2000-05-11
; NUMBER OF SEQ ID NOS: 56
```

```
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3372
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (155)..(3229)
US-10-275-822A-1

Query Match      99.2%; Score 3338.8; DB 16; Length 3372;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 3343; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 15 GAGCCCTGTGGACAAACCTCGTCATTGTTCAGGCACAGAGCGGTAGACCCCTGCTTCTCTAAG 74
Db 1 GGGCCCTGTGGACAAACCTCGTCATTGTTCAGGCACAGAGCGGTAGACCCCTGCTTCTCTAAG 60

QY 75 TGGGCAGCGGACAGCGGCACGCACATTTTCACCTGTCCCGCAGACAAACAGCACCATCTGCT 134
Db 61 TGGGCAGCGGACAGCGGCACGCACATTTTCACCTGTCCCGCAGACAAACAGCACCATCTGCT 120

QY 135 TGGGAGAACCCCTCTCCCTTCTCTGAGAAAGAAAAGATGTGAAATGGGTATTTCCACAGACGA 194
Db 121 TGGGAGAACCCCTCTCCCTTCTCTGAGAAAGAAAAGATGTGAAATGGGTATTTCCACAGACGA 180

QY 195 GAATTTCCGCTATCTCATCTCGTGCTTCAGGGCCAGGGTGAATGTACATCCAGGTGGA 254
Db 181 GAATTTCCGCTATCTCATCTCGTGCTTCAGGGCCAGGGTGAATGTACATCCAGGTGGA 240

QY 255 GCCTGTGCTGGACTACCTGACCTTTCTGCTGCAGAGGTTGAAGGAGCAGATTTCAGAGGAC 314
Db 241 GCCTGTGCTGGACTACCTGACCTTTCTGCTGCAGAGGTTGAAGGAGCAGATTTCAGAGGAC 300

QY 315 AGTCGCCACCTCCGGGAAACATGCAGGCAGTTGAACTGCTGCTGAGCACCTTTGGAGAAGGG 374
Db 301 AGTCGCCACCTCCGGGAAACATGCAGGCAGTTGAACTGCTGCTGAGCACCTTTGGAGAAGGG 360

QY 375 AGTCGGCACCTTGGTTGGACTCGGGAAATTCGTGGAGGCCCTCCGGAGAACCCGGCAGCCC 434
Db 361 AGTCGGCACCTTGGTTGGACTCGGGAAATTCGTGGAGGCCCTCCGGAGAACCCGGCAGCCC 420

QY 435 TCTGGCCGCCCGCTACATGAACCCCTGAGCTCAGGAATTGCCCCCTCTCCATCGTTTGAGAA 494
Db 421 TCTGGCCGCCCGCTACATGAACCCCTGAGCTCAGGAATTGCCCCCTCTCCATCGTTTGAGAA 480

QY 495 CGCTCATGATGAATATCTCAAACCTGCTGAACCTCCTTCAGCCCACCTCTGGTGGACAAGCT 554
Db 481 CGCTCATGATGAATATCTCAAACCTGCTGAACCTCCTTCAGCCCACCTCTGGTGGACAAGCT 540

QY 555 TCTAGTTAGAGACGTCCTTGGATAAGTGCATGGAGAGGAACTGTTGACAATTTGAAGACAG 614
Db 541 TCTAGTTAGAGACGTCCTTGGATAAGTGCATGGAGAGGAACTGTTGACAATTTGAAGACAG 600

QY 615 AAACCGGATTGCTGCTGCAGAAAACAATGGAATGAATCAGGTGTAAGAGAGCTACTAAA 674
Db 601 AAACCGGATTGCTGCTGCAGAAAACAATGGAATGAATCAGGTGTAAGAGAGCTACTAAA 660

QY 675 AAGGATTGTGCAGAAAAGAAACTGGTTCTCTGCAATTTCTGAATGTTCTTCGTCAAACAGG 734
Db 661 AAGGATTGTGCAGAAAAGAAACTGGTTCTCTGCAATTTCTGAATGTTCTTCGTCAAACAGG 720

QY 735 AAACAATGAACCTTGTCCAAAGATTAAACAGGCTCTGATTGCTCAGAAAAGCAATGCAGAGAT 794
Db 721 AAACAATGAACCTTGTCCAAAGATTAAACAGGCTCTGATTGCTCAGAAAAGCAATGCAGAGAT 780

QY 795 TGAGAAATTTATCACAAGTTTGAATGGTCTCAAGTGGAGAGCAACTTCTTTCAACCACAGT 854
Db 781 TGAGAAATTTATCACAAGTTTGAATGGTCTCAAGTGGAGAGCAACTTCTTTCAACCACAGT 840

QY 855 TCAGCCAAATCTGGAGAAGGAGGTCTGGGGCATGGAGAATAAATCATCAGAAATCATCTTT 914
Db 841 TCAGCCAAATCTGGAGAAGGAGGTCTGGGGCATGGAGAATAAATCATCAGAAATCATCTTT 900
```

QY 915 TGCAGATTCTTCTGTAGTTTCAGAAATCAGACACAAGTTTGGCAGAAGGAAGTGTGAGCTG 974
Db 901 TGCAGATTCTTCTGTAGTTTCAGAAATCAGACACAAGTTTGGCAGAAGGAAGTGTGAGCTG 960
QY 975 CTTAGATGAAAGTCTTGGACATAACAGCAACATGGCAGTGAATTCAGGCACCATGGGAAG 1034
Db 961 CTTAGATGAAAGTCTTGGACATAACAGCAACATGGCAGTGAATTCAGGCACCATGGGAAG 1020
QY 1035 TGATTGAGATGAAGAGAATGTGGCAGCAAGAGCATCCCCGGAGCCAGAACTCCAGCTCAG 1094
Db 1021 TGATTGAGATGAAGAGAATGTGGCAGCAAGAGCATCCCCGGAGCCAGAACTCCAGCTCAG 1080
QY 1095 GCCTTACCAAATGGAAGTTGCCAGCCAGCCTTGGAAAGGGAAGAAATATCATCATCTGCCCT 1154
Db 1081 GCCTTACCAAATGGAAGTTGCCAGCCAGCCTTGGAAAGGGAAGAAATATCATCATCTGCCCT 1140
QY 1155 CCCTACAGGGAGTGGAAAAACCCAGAGTGGTGTTTACATTTGCCAAGGATCACTTAGACAA 1214
Db 1141 CCCTACAGGGAGTGGAAAAACCCAGAGTGGTGTTTACATTTGCCAAGGATCACTTAGACAA 1200
QY 1215 GAAGAAAAAAGCATCTGAGCCTGGAAAAAGTTATAGTTCTTGTCAATAAGGTACTGCTAGT 1274
Db 1201 GAAGAAAAAAGCATCTGAGCCTGGAAAAAGTTATAGTTCTTGTCAATAAGGTACTGCTAGT 1260
QY 1275 TGAACAGCTCTTCCGCAAGGAGTTCCAACCATTTTGAAGAAATGGTATCGTGTATTGG 1334
Db 1261 TGAACAGCTCTTCCGCAAGGAGTTCCAACCATTTTGAAGAAATGGTATCGTGTATTGG 1320
QY 1335 ATTAAGTGGTGATACCCAACTGAAAAATATCATTTCCAGAAAGTTGTCAAGTCTGTGATAT 1394
Db 1321 ATTAAGTGGTGATACCCAACTGAAAAATATCATTTCCAGAAAGTTGTCAAGTCTGTGATAT 1380
QY 1395 TATTATCAGTACAGCTCAAATCCTTGA AAACTCCCTCTTAAACTTTGGA AAAATGGAGAAGA 1454
Db 1381 TATTATCAGTACAGCTCAAATCCTTGA AAACTCCCTCTTAAACTTTGGA AAAATGGAGAAGA 1440
QY 1455 TGTGGTGTTC AATTGT CAGACTTTTCCCTCATTATCATTTGATGAATGT CATCACACAA 1514
Db 1441 TGTGGTGTTC AATTGT CAGACTTTTCCCTCATTATCATTTGATGAATGT CATCACACAA 1500
QY 1515 CAAAGAAGCAGTGTATAATAACATCATGAGGCATTATTTGATGCAGAAGTTGAAAAACAA 1574
Db 1501 CAAAGAAGCAGTGTATAATAACATCATGAGGCATTATTTGATGCAGAAGTTGAAAAACAA 1560
QY 1575 TAGACTCAAGAAAGAAAAACAAACCAGTGAATCCCTTCCCTCAGATACTGGGACTAACAGC 1634
Db 1561 TAGACTCAAGAAAGAAAAACAAACCAGTGAATCCCTTCCCTCAGATACTGGGACTAACAGC 1620
QY 1635 TTCACCTGGTGTGGAGGGGCCACGAAGCAAGCCAAAGCTGAAGAACACATTTTAAAACT 1694
Db 1621 TTCACCTGGTGTGGAGGGGCCACGAAGCAAGCCAAAGCTGAAGAACACATTTTAAAACT 1680
QY 1695 ATGTGCCAATCTTGATGCATTTACTATTAAAACTGTTAAAGAAAACTTTGATCAACTGAA 1754
Db 1681 ATGTGCCAATCTTGATGCATTTACTATTAAAACTGTTAAAGAAAACTTTGATCAACTGAA 1740
QY 1755 AAACCAAATACAGGAGCCATGCAAGAAAGTTTGCCATTGCAGATGCAACCAGAGAAGATCC 1814
Db 1741 AAACCAAATACAGGAGCCATGCAAGAAAGTTTGCCATTGCAGATGCAACCAGAGAAGATCC 1800
QY 1815 ATTTAAAGAGAAACTTCTAGAAAATTAATGACAAGGATTCAAACTTATTGTCAAATGAGTCC 1874
Db 1801 ATTTAAAGAGAAACTTCTAGAAAATTAATGACAAGGATTCAAACTTATTGTCAAATGAGTCC 1860
QY 1875 AATGTCAGATTTTGGAACTCAACCCCTATGAACAAATGGGCCATTCAAATGGA AAAAAGC 1934
Db 1861 AATGTCAGATTTTGGAACTCAACCCCTATGAACAAATGGGCCATTCAAATGGA AAAAAGC 1920
QY 1935 TGCAAAAAAAGGAAATCGCAAGAACGTGTTTGTGCAGAACATTTGAGGAAGTACAATGA 1994
Db 1921 TGCAAAAAAAGGAAATCGCAAGAACGTGTTTGTGCAGAACATTTGAGGAAGTACAATAA 1980

QY 1995 GGCCCTACAAATTAATGACACAATTCGAATGATAGATGCGTATACTCATCTTTGAAACTTT 2054
Db 1981 GGCCCTACAAATTAATGACACAATTCGAATGATAGATGCGTATACTCATCTTTGAAACTTT 2040
QY 2055 CTATAATGAAGAGAAAAAGATAAGAAAGTTTGCAGTCTATAGAAAGATGATAGTATGAGGGTGG 2114
Db 2041 CTATAATGAAGAGAAAAAGATAAGAAAGTTTGCAGTCTATAGAAAGATGATAGTATGAGGGTGG 2100
QY 2115 TGATGATGAGTATTGTGATGGTGATGAAGATGAGGATGATTTTAAAGAAAACTTTTGA AACT 2174
Db 2101 TGATGATGAGTATTGTGATGGTGATGAAGATGAGGATGATTTTAAAGAAAACTTTTGA AACT 2160
QY 2175 GGATGAAACAGATAGATTTTCTCATGACTTTTATTTTGTGAAAACAATAAAATGTTGAAAAG 2234
Db 2161 GGATGAAACAGATAGATTTTCTCATGACTTTTATTTTGTGAAAACAATAAAATGTTGAAAAG 2220
QY 2235 GCTGGCTGAAAAACCCAGAAATATGAAATGAAAAAGCTGACCAAAATTAAGAAATACCATAAAT 2294
Db 2221 GCTGGCTGAAAAACCCAGAAATATGAAATGAAAAAGCTGACCAAAATTAAGAAATACCATAAAT 2280
QY 2295 GGAGCAATATACTAGGACTGAGGAATCAGCACAGGAGTAATCTTTACAAAAACACGACA 2354
Db 2281 GGAGCAATATACTAGGACTGAGGAATCAGCACAGGAGTAATCTTTACAAAAACACGACA 2340
QY 2355 GAGTGCATATGCGCTTTCCAGTGGATTACTGAAAAATGAAAAATTTGCTGAAGTAGGAGT 2414
Db 2341 GAGTGCATATGCGCTTTCCAGTGGATTACTGAAAAATGAAAAATTTGCTGAAGTAGGAGT 2400
QY 2415 CAAAGCCCCACCATCTGATTGGAGCTGGACACAGCAGTGAAGTTCAAAACCCATGACACAGAA 2474
Db 2401 CAAAGCCCCACCATCTGATTGGAGCTGGACACAGCAGTGAAGTTCAAAACCCATGACACAGAA 2460
QY 2475 TGAACAAAAAGAAAGTCATTAGTAAATTTCCGACTGGAAAAATCAATCTGCTTATCGCTAC 2534
Db 2461 TGAACAAAAAGAAAGTCATTAGTAAATTTCCGACTGGAAAAATCAATCTGCTTATCGCTAC 2520
QY 2535 CACAGTGGCAGAAAGAGTCTGGATATTAAAGAAATGTAACATTTGTTATCCGTTATGGTCT 2594
Db 2521 CACAGTGGCAGAAAGAGTCTGGATATTAAAGAAATGTAACATTTGTTATCCGTTATGGTCT 2580
QY 2595 CGTACCCAATGAAATAGCCATGGTCCAGGCCCGTGGTCCAGCCAGAGCTGATGAGAGCAC 2654
Db 2581 CGTACCCAATGAAATAGCCATGGTCCAGGCCCGTGGTCCAGCCAGAGCTGATGAGAGCAC 2640
QY 2655 CTACGTCCTGGTTGCTCACAGTGGTTTCCAGGATTATCGAAACATGAGACAGTTAATGATTT 2714
Db 2641 CTACGTCCTGGTTGCTCACAGTGGTTTCCAGGATTATCGAAACATGAGACAGTTAATGATTT 2700
QY 2715 CCGAGAGAAAGATGATGTATAAAGCTATACATTTGTGTTCAAAATATGAAACCCAGAGGAGTA 2774
Db 2701 CCGAGAGAAAGATGATGTATAAAGCTATACATTTGTGTTCAAAATATGAAACCCAGAGGAGTA 2760
QY 2775 TGCTCATAAAGATTTTGGAAATTACAGATGCAAAAGTATAATGGAAAAAGAAAAATGAAACCAA 2834
Db 2761 TGCTCATAAAGATTTTGGAAATTACAGATGCAAAAGTATAATGGAAAAAGAAAAATGAAACCAA 2820
QY 2835 GAGAAATATTGCCAAGCATTACAAGAAATAACCCATCACTAACTTTTCCCTTTGCAAAAA 2894
Db 2821 GAGAAATATTGCCAAGCATTACAAGAAATAACCCATCACTAACTTTTCCCTTTGCAAAAA 2880
QY 2895 CTGCAGTGTCTAGCCTGTTCTGGGGAAGATATCCATGTAAATTGAGAAAAATGCATCACGT 2954
Db 2881 CTGCAGTGTCTAGCCTGTTCTGGGGAAGATATCCATGTAAATTGAGAAAAATGCATCACGT 2940
QY 2955 CAATATGACCCCAAGAAATTCAGGAACTTTTACATTTGTAAAGAGAAAACAAAGCATCTGCAAAA 3014
Db 2941 CAATATGACCCCAAGAAATTCAGGAACTTTTACATTTGTAAAGAGAAAACAAAGCATCTGCAAAA 3000
QY 3015 GAAGTGTCCGACTATCAAAATAAATGGTGAATCATCTGCAAAATGTGGCCAGGCTTGGGG 3074
Db 3001 GAAGTGTCCGACTATCAAAATAAATGGTGAATCATCTGCAAAATGTGGCCAGGCTTGGGG 3060
QY 3075 AACAAATGATGGTGCAAAAAGGCTTAGATTTTGCCCTTGTCTCAAAAATAAGGAAATTTTGTGATG 3134

Db 3061 AACAATGATGGTGCACAAAGGCTTAGATTTGCCCTTGCTCAAAATAAGGAATTTTGTAGT 3120
Qy 3135 GGTTCCTCAAAATAAATTCACAAAGAAACAATACAAAAGTGGGTAGAAATFACCTATCAC 3194
Db 3121 GGTTCCTCAAAATAAATTCACAAAGAAACAATACAAAAGTGGGTAGAAATFACCTATCAC 3180
Qy 3195 ATTTCCCAATCTTGACTATTTCAGAAATGCTGTTTATTTAGTGATGAGGATAGCACTTGAT 3254
Db 3181 ATTTCCCAATCTTGACTATTTCAGAAATGCTGTTTATTTAGTGATGAGGATAGCACTTGAT 3240
Qy 3255 TGAAGATCTTTTAAATACTATCAGTTAAACATTTTAATATGATTATGATTAAATGATTTC 3314
Db 3241 TGAAGATCTTTTAAATACTATCAGTTAAACATTTTAATATGATTATGATTAAATGATTTC 3300
Qy 3315 ATTATGCTACAGAACTGACATAAGAAATCAATAAAATGATTGTTTACTCT 3364
Db 3301 ATTATGCTACAGAACTGACATAAGAAATCAATAAAATGATTGTTTACTCT 3350

RESULT 9

US-10-094-749-828
; Sequence 828, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHIKO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 828
; LENGTH: 1776
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-094-749-828

Query Match 52.5%; Score 1766; DB 15; Length 1776;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1769; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 1592 ACAAAACAGTGATCCCTTCCTCAGATACTGGGACTAACAGCTTACCTGGTGGAG 1651
Db 2 ACAAAACAGTGATCCCTTCCTCAGATACTGGGACTAACAGCTTACCTGGTGGAG 61
Qy 1652 GGGCCACGAAGCAAGCCAAAGCTGAAGAACACATTTTAAACATATGTGCCAATCTTGATG 1711
Db 62 GGGCCACGAAGCAAGCCAAAGCTGAAGAACACATTTTAAACATATGTGCCAATCTTGATG 121
Qy 1712 CATTTACTATTAAACTGTTAAAGAAAACCTTGATCAACTGAAAAACCAATACAGGAGC 1771

Db 122 CATTTACTATTAAAACTGTTAAAGAAAACCTTGATCAACTGAAAAACCAATACAGGAGC 181
Qy 1772 CATGCAAGAAGTTTGCCATTGCAGATGCAACACAGAGAAGATCCATTTTAAAGAGAAACTTC 1831
Db 182 CATGCAAGAAGTTTGCCATTGCAGATGCAACACAGAGAAGATCCATTTTAAAGAGAAACTTC 241
Qy 1832 TAGAAATAATGACAAGGATTCAAACTTTATGTCAAATGAGTCCCAATGTCAAGATTTTGGAA 1891
Db 242 TAGAAATAATGACAAGGATTCAAACTTTATGTCAAATGAGTCCCAATGTCAAGATTTTGGAA 301
Qy 1892 CTCAAACCTTATGAACAATGGGCCATTCAATATGGAAGAAAAAGCTGCAAAAAAGGAAATC 1951
Db 302 CTCAAACCTTATGAACAATGGGCCATTCAATATGGAAGAAAAAGCTGCAAAAAAGGAAATC 361
Qy 1952 GCAAAAGAACGTTTGTGCAGAACATTTGAGGAAGTACAATGAGGCCCTACAAATTAATG 2011
Db 362 GCAAAAGAACGTTTGTGCAGAACATTTGAGGAAGTACAATGAGGCCCTACAAATTAATG 421
Qy 2012 ACACAAATTCGAATGATAGATGCGTATATCTCATCTTGAACCTTTCTATAATGAAGAGAAAG 2071
Db 422 ACACAAATTCGAATGATAGATGCGTATATCTCATCTTGAACCTTTCTATAATGAAGAGAAAG 481
Qy 2072 ATAAGAAAGTTTGCAGTCATAGAGATGATAGTATGAGGGTGGTGATGATGAGTATTGTG 2131
Db 482 ATAAGAAAGTTTGCAGTCATAGAGATGATAGTATGAGGGTGGTGATGATGAGTATTGTG 541
Qy 2132 ATGTTGATGAAGATGAGGATGATTAAAGAAAACCTTTGAAACTGGATGAAACAGATAGAT 2191
Db 542 ATGTTGATGAAGATGAGGATGATTAAAGAAAACCTTTGAAACTGGATGAAACAGATAGAT 601
Qy 2192 TTCTCATGACTTTTATTTTGAACCAATAAATGTTGAAAAGGCTGCTGAAAAACCCAG 2251
Db 602 TTCTCATGACTTTTATTTTGAACCAATAAATGTTGAAAAGGCTGCTGAAAAACCCAG 661
Qy 2252 AATATGAAAAATGAAAAGCTGACCAAAATTAAGAAATACCATAATGGAGCAATATACTAGGA 2311
Db 662 AATATGAAAAATGAAAAGCTGACCAAAATTAAGAAATACCATAATGGAGCAATATACTAGGA 721
Qy 2312 CTGAGGAATCAGCACGAGGAATAATCTTTACAAAAACACGACAGAGTGCATATCGGCTTT 2371
Db 722 CTGAGGAATCAGCACGAGGAATAATCTTTACAAAAACACGACAGAGTGCATATCGGCTTT 781
Qy 2372 CCCAGTGGATTACTGAAAATGAAAAATTTGCTGAAAGTAGGAGTCAAAGCCCACCATCTGA 2431
Db 782 CCCAGTGGATTACTGAAAATGAAAAATTTGCTGAAAGTAGGAGTCAAAGCCCACCATCTGA 841
Qy 2432 TTGGAGCTGGACACAGCAGTGAAGTTCAAACCCCATGACACAGAAATGAACAAAAAGAGTCA 2491
Db 842 TTGGAGCTGGACACAGCAGTGAAGTTCAAACCCCATGACACAGAAATGAACAAAAAGAGTCA 901
Qy 2492 TTAGTAAATTTGCGACTGGAAAAATCAATCTGCTTATCGCTACCACAGTGGCAGAGAAG 2551
Db 902 TTAGTAAATTTGCGACTGGAGAATAAATCTGCTTATCGCTACCACAGTGGCAGAGAAG 961
Qy 2552 GTCTGGATATTAAAGAAATGTAACATTGTTATCCGTTATGGTCTCGTCACCAATGAAATAG 2611
Db 962 GTCTGGATATTAAAGAAATGTAACATTGTTATCCGTTATGGTCTCGTCACCAATGAAATAG 1021
Qy 2612 CCATGGTCCAGGCCCGTGGTCGAGCCAGAGCTGATGAGAGCACCTACGTCCTGTTGCTC 2671
Db 1022 CCATGGTCCAGGCCCGTGGTCGAGCCAGAGCTGATGAGAGCACCTACGTCCTGTTGCTC 1081
Qy 2672 ACAGTGGTTCAGGAGTTATCGAAATGAGACAGTGAATGATTTCCGAGAGAAGATGATGT 2731
Db 1082 ACAGTGGTTCAGGAGTTATCGAAATGAGACAGTGAATGATTTCCGAGAGAAGATGATGT 1141
Qy 2732 ATAAAGCTATACATTTGTTCAAAAATATGAAACCCAGAGGAGTATGCTCAFAAGATTTTGG 2791
Db 1142 ATAAAGCTATACATTTGTTCAAAAATATGAAACCCAGAGGAGTATGCTCAFAAGATTTTGG 1201
Qy 2792 AATTACAGATGCAAAAGTATAATGGAAGAAAAAGAAATGAAAAACCAAGAGAAATATTGCCAAGC 2851
Db 1202 AATTACAGATGCAAAAGTATAATGGAAGAAAAAGAAATGAAAAACCAAGAGAAATATTGCCAAGC 1261


```
QY 2852 ATTACAAGAATAACCCATCACTAATAACTTTCCTTTGCAAAAACGTCAGTGTGCTAGCCT 2911
Db 1262 ATTACAAGAATAACCCATCACTAATAACTTTCCTTTGCAAAAACGTCAGTGTGCTAGCCT 1321
QY 2912 GTTCTGGGAAGATATCCATGTAATTGAGAAAATGCATCACGTCATATGACCCCGAAT 2971
Db 1322 GTTCTGGGAAGATATCCATGTAATTGAGAAAATGCATCACGTCATATGACCCCGAAT 1381
QY 2972 TCAAGGAACCTTTACATTTGTAAGAGAAAAACAAGCAGCTGCAAAAGAGTGTGCCGACTATC 3031
Db 1382 TCAAGGAACCTTTACATTTGTAAGAGAAAAACAAGCAGCTGCAAAAGAGTGTGCCGACTATC 1441
QY 3032 AAATAAATGGTGAATCATCTGCAAAATGTGGCCAGGCTTGGGAAACAATGATGGTGACA 3091
Db 1442 AAATAAATGGTGAATCATCTGCAAAATGTGGCCAGGCTTGGGAAACAATGATGGTGACA 1501
QY 3092 AAGGCTTAGATTTGCCTTGTCTCAAAATAAGGAATTTTGTAGTGGTTTCAAAAATAATT 3151
Db 1502 AAGGCTTAGATTTGCCTTGTCTCAAAATAAGGAATTTTGTAGTGGTTTCAAAAATAATT 1561
QY 3152 CAACAAAGAAACAATACAAAAAGTGGGTAGAATTACCTATCACATTTCCCAATCTTGACT 3211
Db 1562 CAACAAAGAAACAATACAAAAAGTGGGTAGAATTACCTATCACATTTCCCAATCTTGACT 1621
QY 3212 ATTCAGAATGCTGTTTATTATTAGTAGGAGGATTAGCACCTTGATTGAAGATTCTTTTAAA 3271
Db 1622 ATTCAGAATGCTGTTTATTATTAGTAGGAGGATTAGCACCTTGATTGAAGATTCTTTTAAA 1681
QY 3272 TACTATCAGTTAAACATTTAATATGATTATGATTAATGATTATTCATTTACAGAACTG 3331
Db 1682 TACTATCAGTTAAACATTTAATATGATTATGATTAATGATTATTCATTTACAGAACTG 1741
QY 3332 ACATAAGAATCAATAAATGATTGTTTACTCTG 3365
Db 1742 ACATAAGAATCAATAAATGATTGTTTACTCTG 1775

RESULT 10
US-10-275-822A-10
; Sequence 10, Application US/10275822A
; Publication No. US20040086500A1
; GENERAL INFORMATION:
; APPLICANT: Bahr, Georges
; APPLICANT: Cocude, Cecile
; APPLICANT: Capron, Andre
; TITLE OF INVENTION: RH116 Polypeptide and its Fragments and Polynucleotides Encoding
; TITLE OF INVENTION: said Polypeptides and Therapeutic Uses
; FILE REFERENCE: 017753-170
; CURRENT APPLICATION NUMBER: US/10/275,822A
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: PCT/FR01/01441
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: FR 00/06,030
; PRIOR FILING DATE: 2000-05-11
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 1443
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 927
; OTHER INFORMATION: n = A,T,C or G
US-10-275-822A-10

Query Match 42.7%; Score 1435.6; DB 16; Length 1443;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1438; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 687 GAAAGAAAACGGTTCTCTGCAATTTCTGAATGTTCTTCGTCAACAGGAACAATGAACT 746
|||||
```

```
Db 1 GAAAGAAAACGGTTCTCTGCAATTTCTGAATGTTCTTCGTCAACAGGAACAATGAACT 60
QY 747 TGTCCAAGAGTTAAACAGGCTCTGATTGCTCAGAAAAGCAATGCAGAGATTGAGAAATTTATC 806
Db 61 TGTCCAAGAGTTAAACAGGCTCTGATTGCTCAGAAAAGCAATGCAGAGATTGAGAAATTTATC 120
QY 807 ACAAGTTGATGTCCTCAAGTGGAGAGCAAACTTCTTTCAACCACAGATTCAGCCAAATCT 866
Db 121 ACAAGTTGATGTCCTCAAGTGGAGAGCAAACTTCTTTCAACCACAGATTCAGCCAAATCT 180
QY 867 GGAGAAAGGAGTCTGGGGCATGGAGAAATAACTCATCAGAAATCATCTTTTGCAGATTCCTC 926
Db 181 GGAGAAAGGAGTCTGGGGCATGGAGAAATAACTCATCAGAAATCATCTTTTGCAGATTCCTC 240
QY 927 TGTAGTTTTCAGAAATCAGACACAAAGTTTGGCAGAAAGAGTGTGAGTGTGATGAAAG 986
Db 241 TGTAGTTTTCAGAAATCAGACACAAAGTTTGGCAGAAAGAGTGTGAGTGTGATGAAAG 300
QY 987 TCTTGGACATAACAGCAACATGGGCAAGTATTCAGGCACCATGGGAAAGTATTCAGATGA 1046
Db 301 TCTTGGACATAACAGCAACATGGGCAAGTATTCAGGCACCATGGGAAAGTATTCAGATGA 360
QY 1047 AGAGAAATGTGCAGCAAGAGCATCCCGGAGCCAGAACTCCAGCTCAGGCCCTTACCAGAT 1106
Db 361 AGAGAAATGTGCAGCAAGAGCATCCCGGAGCCAGAACTCCAGCTCAGGCCCTTACCAGAT 420
QY 1107 GGAAGTTGCCAGCCAGCCCTTGGAGGGAGAAATATCATCATCTGCTCCCTACAGGGAG 1166
Db 421 GGAAGTTGCCAGCCAGCCCTTGGAGGGAGAAATATCATCATCTGCTCCCTACAGGGAG 480
QY 1167 TGGAAAAACCCAGAGTGGCTGTTTACATTGCCAAGGATCACTTAGACAAGAAAAAAGC 1226
Db 481 TGGAAAAACCCAGAGTGGCTGTTTACATTGCCAAGGATCACTTAGACAAGAAAAAAGC 540
QY 1227 ATCTGAGCCTGGAAAAAGTTATAGTTCCTTGTCAATAAGGTACTGCTAGTTGAACAGCTCT 1286
Db 541 ATCTGAGCCTGGAAAAAGTTATAGTTCCTTGTCAATAAGGTACTGCTAGTTGAACAGCTCT 600
QY 1287 CCGCAAGGAGTTCCAACCATTTTGAAGAAAATGGTATCGTGTATTGGATTAAAGTGTGA 1346
Db 601 CCGCAAGGAGTTCCAACCATTTTGAAGAAAATGGTATCGTGTATTGGATTAAAGTGTGA 660
QY 1347 TACCCAACTGAAAAATATCATTTCCAGAAAGTTGTCAAGTCCCTGTGATATTATTATCAGTAC 1406
Db 661 TACCCAACTGAAAAATATCATTTCCAGAAAGTTGTCAAGTCCCTGTGATATTATTATCAGTAC 720
QY 1407 AGCTCAAACTCTTGAAAACTCCCTCTTAAACTTGGAAAATGGAGAAAGATGCTGGTGTCA 1466
Db 721 AGCTCAAACTCTTGAAAACTCCCTCTTAAACTTGGAAAATGGAGAAAGATGCTGGTGTCA 780
QY 1467 ATTGTCAGACTTTTCCCTCATTTATCATTTGATGAATGTATCACACCAACAAAGAGCAGT 1526
Db 781 ATTGTCAGACTTTTCCCTCATTTATCATTTGATGAATGTATCATCACACCAACAAAGAGCAGT 840
QY 1527 GTATAATAACATCATGAGGCAATTTTGTATGCAGAAAGTTGAAAAACAATAGACTCAAGAA 1586
Db 841 GTATAATAACATCATGAGGCAATTTTGTATGCAGAAAGTTGAAAAACAATAGACTCAAGAA 900
QY 1587 AGAAAAACAACCCAGTGTATCCCTTCCCTCAGATACCTGGGACTAACAGCTTCACCTGGTGT 1646
Db 901 AGAAAAACAACCCAGTGTATCCCTTCCCTCAGATACCTGGGACTAACAGCTTCACCTGGTGT 960
QY 1647 TGGAGGGGCCAGGAAGCAAGCCAAAGCTGAAGAACACATTTTAAACTATGTGCCAATCT 1706
Db 961 TGGAGGGGCCAGGAAGCAAGCCAAAGCTGAAGAACACATTTTAAACTATGTGCCAATCT 1020
QY 1707 TGATGCATTTACTATTAAAACTGTTAAAGAAAACTTGTATCAACTGAAAAACCAAATACA 1766
Db 1021 TGATGCATTTACTATTAAAACTGTTAAAGAAAACTTGTATCAACTGAAAAACCAAATACA 1080
QY 1767 GGAGCCATGCAAGAAAGTTTGGCCATTGCAGATGCAACCCAGAGAAAGATCCATTTTAAAGAGAA 1826
Db 1081 GGAGCCATGCAAGAAAGTTTGGCCATTGCAGATGCAACCCAGAGAAAGATCCATTTTAAAGAGAA 1140
```


QY 2996 AAAACAAAGCACTGCAAAAAGAAGTGTCCGACTATCAATAAATGCTGAAATCATCTGCA 3055
Db 1021 AAAACAAAGCACTGCAAAAAGAAGTGTCCGACTATCAATAAATGCTGAAATCATCTGCA 1080
QY 3056 AATGTGCCAGGCTTTGGGGAACAATGATGGTGCAAAAGCTTAGATTGGCCTTGCTCTCA 3115
Db 1081 AATGTGCCAGGCTTTGGGGAACAATGATGGTGCAAAAGCTTAGATTGGCCTTGCTCTCA 1140
QY 3116 AAATAAGGAATTTTGTAGTGGTTTTTCAAAAATAATTCAACAAAGAAACAATACAAAAGT 3175
Db 1141 AAATAAGGAATTTTGTAGTGGTTTTTCAAAAATAATTCAACAAAGAAACAATACAAAAGT 1200
QY 3176 GGGTAGAATTACCTATCACATTTCCCAATCTTGACTATTTCAGAAATGCTGTTATTAGTG 3235
Db 1201 GGGTAGAATTACCTATCACATTTCCCAATCTTGACTATTTCAGAAATGCTGTTATTAGTG 1260
QY 3236 ATGAGGATTAGCACTTGATTGAAGATTCTTTTAAATACTATCAGTTAAACATTTTAATAT 3295
Db 1261 ATGAGGATTAGCACTTGATTGAAGATTCTTTTAAATACTATCAGTTAAACATT--AATA 1318
QY 3296 GATTATGATTAATGATTTCATTTATGCTACAGAACTGACATAAGAATC---AATAAAATGA 3352
Db 1319 TGTATTGATTAAATGATTTCATTTATGCTACAGAACTGACAAATNGAGNCTCAATAAAATGA 1378
QY 3353 TTGTTTTACTCTG 3365
Db 1379 TTGTTTTACTCTG 1391

RESULT 12
US-10-275-822A-7
; Sequence 7, Application US/10275822A
; Publication No. US20040086500A1
; GENERAL INFORMATION:
; APPLICANT: Bahr, Georges
; APPLICANT: Cocude, Cecile
; APPLICANT: Capron, Andre
; TITLE OF INVENTION: RH16 Polypeptide and its Fragments and Polynucleotides Encoding
; TITLE OF INVENTION: said Polypeptides and Therapeutic Uses
; FILE REFERENCE: 01753-170
; CURRENT APPLICATION NUMBER: US/10/275,822A
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: PCT/FR01/01441
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: FR 00/06,030
; PRIOR FILING DATE: 2000-05-11
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 1284
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1261
; OTHER INFORMATION: n = A,T,C or G
US-10-275-822A-7

Query Match 37.3%; Score 1254.6; DB 16; Length 1284;
Best Local Similarity 99.6%; Pred. No. 5.1e-283;
Matches 1257; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 2103 TGATGAGGGTGGTGATGATGAGTATTGTGATGGTGATGAAGATGAGGATGATTTAAAGAA 2162
Db 1 TGATGAGGGTGGTGATGATGAGTATTGTGATGGTGATGAAGATGAGGATGATTTAAAGAA 60
QY 2163 ACCTTTGAAACTGGATGAAACAGATAGATTTCTCATGACTTTTATTTTGTGAAACAAATAA 2222
Db 61 ACCTTTGAAACTGGATGAAACAGATAGATTTCTCATGACTTTTATTTTGTGAAACAAATAA 120
QY 2223 AATGTTGAAAAGGCTGGCTGAAAACCCAGAAATATGAAAATGAAAAGCTGACCAAAATTAAG 2282

Db 121 AATGTTGAAAAGGCTGGCTGAAAACCCAGAAATATGAAAATGAAAAGCTGACCAAAATTAAG 180
QY 2283 AAATACCATAAATGGAGCAATATATACTAGGACTGAGGAATCAGCAGAGGAATAAATCTTTAC 2342
Db 181 AAATACCATAAATGGAGCAATATATACTAGGACTGAGGAATCAGCAGAGGAATAAATCTTTAC 240
QY 2343 AAAAACACGACAGAGTGCCATATGCGCTTTCCAGTGGATTACTGAAAAATGAAAAATTTGC 2402
Db 241 AAAAACACGACAGAGTGCCATATGCGCTTTCCAGTGGATTACTGAAAAATGAAAAATTTGC 300
QY 2403 TGAAGTAGGAGTCAAAGCCCACCACCTCTGATTGGAGCTGGACACAGCAGTGAAGTTCAAACC 2462
Db 301 TGAAGTAGGAGTCAAAGCCCACCACCTCTGATTGGAGCTGGACACAGCAGTGAAGTTCAAACC 360
QY 2463 CATGACACAGAAATGAACAAAGAAAGTCAATAGTAAATTTCCGACTGGAAAAATCAATCT 2522
Db 361 CATGACACAGAAATGAACAAAGAAAGTCAATAGTAAATTTCCGACTGGAAAAATCAATCT 420
QY 2523 GCTTATCGTTACCACAGTGGCAGAGAAAGGTCTGGATATTAAAGAAATGTAACATTTGTTAT 2582
Db 421 GCTTATCGTTACCACAGTGGCAGAGAAAGGTCTGGATATTAAAGAAATGTAACATTTGTTAT 480
QY 2583 CCGTTATGCTCGTCACCAATGAATAGCCATGGTCCAGGCCGTGGTCGAGCCAGAGC 2642
Db 481 CCGTTATGCTCGTCACCAATGAATAGCCATGGTCCAGGCCGTGGTCGAGCCAGAGC 540
QY 2643 TGATGAGAGCACCTACGTCCTGTTGCTCACAGTGGTTCAGGAGTTATCGAACATGAGAC 2702
Db 541 TGATGAGAGCACCTACGTCCTGTTGCTCACAGTGGTTCAGGAGTTATCGAACGTGAGAC 600
QY 2703 AGTTAATGATTTCCGAGAGAGATGATGTATAAAGCTATACATTTGTTCAAATAATGAA 2762
Db 601 AGTTAATGATTTCCGAGAGAGATGATGTATAAAGCTATACATTTGTTCAAATAATGAA 660
QY 2763 ACCAGAGGATGATGCTCATAAGATTTTGGAAATTACAGATGCAAAAGTATAATGGAAGAA 2822
Db 661 ACCAGAGGATGATGCTCATAAGATTTTGGAAATTACAGATGCAAAAGTATAATGGAAGAA 720
QY 2823 AATGAAAACCAAGAGAAATATTGCCAAGCATTACAAGAAATAACCCCATCACTAATAACTTT 2882
Db 721 AATGAAAACCAAGAGAAATATTGCCAAGCATTACAAGAAATAACCCCATCACTAATAACTTT 780
QY 2883 CTTTTCGAAAAAAGTGCAGTGTGCTAGCCTGTTCTGSGGAGAGATATCCATGTAATTGAGAA 2942
Db 781 CTTTTCGAAAAAAGTGCAGTGTGCTAGCCTGTTCTGSGGAGAGATATCCATGTAATTGAGAA 840
QY 2943 AATGCATCAGTCAATATGACCCCAAGATTTCAAGGAACCTTACATTTGTAAGAGAAAAACAA 3002
Db 841 AATGCATCAGTCAATATGACCCCAAGATTTCAAGGAACCTTACATTTGTAAGAGAAAAACAA 900
QY 3003 AGCACTGCAAAAGAAAGTGTGCCGACTATCAAAATAAATGTTGAAATCATCTGCAAAATGGG 3062
Db 901 AACACTGCAAAAGAAAGTGTGCCGACTATCAAAATAAATGTTGAAATCATCTGCAAAATGGG 960
QY 3063 CCAGGCTTGGGAAACAATGATGGTGCAAAAAGGCTTAGATTGCTTGCCTTGTCTCAAAAATAAG 3122
Db 961 CCAGGCTTGGGAAACAATGATGGTGCAAAAAGGCTTAGATTGCTTGCCTTGTCTCAAAAATAAG 1020
QY 3123 GAATTTTGTAGTGGTTTTTCAAAAATAAATTCAACAAAGAAACAATACAAAAGTGGGTAGA 3182
Db 1021 GAATTTTGTAGTGGTTTTTCAAAAATAAATTCAACAAAGAAACAATACAAAAGTGGGTAGA 1080
QY 3183 ATTACCTATCACATTTCCCAATCTTGACTATTTCAGAAATGCTGTTTATTTTAGTGATGAGGA 3242
Db 1081 ATTACCTATCACATTTCCCAATCTTGACTATTTCAGAAATGCTGTTTATTTTAGTGATGAGGA 1140
QY 3243 TTAGCACTTGATTGAAGATTTCTTTTAAATACTATCAGTTAAACATTTAATATGATTATG 3302
Db 1141 TTAGCACTTGATTGAAGATTTCTTTTAAATACTATCAGTTAAACATTTAATATGATTATG 1200
QY 3303 ATTAATGTATTCAATTATGCTACAGAACTGACATAAGAATCAATAAAAATGATTGTTTTACT 3362
Db 1201 ATTAATGTATTCAATTATGCTACAGAACTGACATAAGAATCAATAAAAATGATTGTTTTACC 1260

QY 3363 CT 3364
|
Db 1261 NT 1262

RESULT 13

US-10-106-698-811
; Sequence 811, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: PA005PI
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 811
; LENGTH: 609
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (32)..(32)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (600)..(601)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (606)..(606)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-106-698-811

Query Match 17.2%; Score 577.2; DB 15; Length 609;
Best Local Similarity 98.6%; Pred. No. 1.5e-124;
Matches 573; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

QY 2785 ATTTTGAATTACAGATGCAAGTATAATGAAAAAGAAATGAAAAACCAAGAGAAATATT 2844
|
Db 5 ATTTTGAATTACAGATGCAAGTATAATGAAAAAGAAATGAAAAACCAAGAGAAATATT 64
|
QY 2845 GCCAAGCATTACAAGATAACCCATCACTAATACTTTCTTTGCAAAAACCTGCAGTGTG 2904
|
Db 65 GCCARGCATTMCARGAATAMCCCATCACTAATACTTTCTTTGCAAAAACCTGCAGTGTG 124
|
QY 2905 CTAGCCTGTTCTGGGGAAGATATCCATGTAATTGAGAAAAATGCATCAGTCAATATGACC 2964
|
Db 125 CTAGCCTGTTCTGGGGAAGATATCCATGTAATTGAGAAAAATGCATCAGTCAATATGACC 184
|
QY 2965 CCAGAATTCAAGGAACCTTACATTGTAAGAGAAAAACAAAGCAGTGCAGAAAGAGTGTGCC 3024
|
Db 185 CCAGAATTCAAGGAACCTTACATTGTAAGAGAAAAACAAARCACTGCAAAAAGAGTGTGCC 244
|
QY 3025 GACTATCAAAATAAATGGTGAAATCATCTGCAAAATGTGGCAGGCTTGGGGAACAATGATG 3084
|
Db 245 GACTATCAAAATAAATGGTGAAATCATCTGCAAAATGTGGCAGGCTTGGGGAACAATGATG 304
|
QY 3085 GTGCACAAAGGCTTAGATTTCCTTGTCTCAAAATAAGGAATTTTGTAGTGGTTTTCAAA 3144
|
Db 305 GTGCACAAAGGCTTAGATTTCCTTGTCTCAAAATAAGGAATTTTGTAGTGGTTTTCAAA 364
|
QY 3145 AATAATTCAACAAAGAAACAATACAAAAGTGGGTAGAATTACCTATCACATTTCCTCAAT 3204
|
Db 365 AATAATTCAACAAAGAAACAATACAAAAGTGGGTAGAATTACCTATCACATTTCCTCAAT 424
|
QY 3205 CTTGACTATTACAGATGCTGTTTATTATTAGTATGAGGATTAGCACITGATTGAAGATTCT 3264

Db 425 CTTGACTATTACAGATGCTGTTTATTATTAGTATGAGGATTAGCACTTGAATTGAAGATTCT 484
|
QY 3265 TTTAAAAATACTATCAGTTAAACATTAAATATATGATTATGATTAAATGTTATTATTATGCTAC 3324
|
Db 485 TTTAAAAATACTATCAGTTAAACATTAAATATATGATTATGATTAAATGTTATTATTATGCTAC 544
|
QY 3325 AGAACTGACATAAGAATCAATAAAATGATTGTTTACTCTG 3365
|
Db 545 AGAACTGACATAAGAATCAATAAAATGATTGTTTACTCTG 585
|

RESULT 14

US-10-264-049-69
; Sequence 69, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA133PI
; CURRENT APPLICATION NUMBER: US/10/264,049
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 69
; LENGTH: 609
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (32)..(32)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (600)..(601)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (606)..(606)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-264-049-69

Query Match 17.2%; Score 577.2; DB 16; Length 609;
Best Local Similarity 98.6%; Pred. No. 1.5e-124;
Matches 573; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

QY 2785 ATTTTGAATTACAGATGCAAGTATAATGAAAAAGAAATGAAAAACCAAGAGAAATATT 2844
|
Db 5 ATTTTGAATTACAGATGCAAGTATAATGAAAAAGAAATGAAAAACCAAGAGAAATATT 64
|
QY 2845 GCCAAGCATTACAAGATAACCCATCACTAATACTTTCTTTGCAAAAACCTGCAGTGTG 2904
|
Db 65 GCCARGCATTMCARGAATAMCCCATCACTAATACTTTCTTTGCAAAAACCTGCAGTGTG 124
|
QY 2905 CTAGCCTGTTCTGGGGAAGATATCCATGTAATTGAGAAAAATGCATCAGTCAATATGACC 2964
|
Db 125 CTAGCCTGTTCTGGGGAAGATATCCATGTAATTGAGAAAAATGCATCAGTCAATATGACC 184
|
QY 2965 CCAGAATTCAAGGAACCTTTTACATTGTAAGAGAAAAACAAAGCAGTGCAGAAAGAGTGTGCC 3024
|
Db 185 CCAGAATTCAAGGAACCTTTTACATTGTAAGAGAAAAACAAARCACTGCAAAAAGAGTGTGCC 244
|
QY 3025 GACTATCAAAATAAATGGTGAAATCATCTGCAAAATGTGGCAGGCTTGGGGAACAATGATG 3084
|
Db 245 GACTATCAAAATAAATGGTGAAATCATCTGCAAAATGTGGCAGGCTTGGGGAACAATGATG 304
|
QY 3085 GTGCACAAAGGCTTAGATTTCCTTGTCTCAAAATAAGGAATTTTGTAGTGGTTTTCAAA 3144
|
Db 305 GTGCACAAAGGCTTAGATTTCCTTGTCTCAAAATAAGGAATTTTGTAGTGGTTTTCAAA 364

QY3145AATAATTCAACAAGAAACAATACAAAAAGTGGTAGAATTACCTATCACATTTCCCAAT3204

Db|||||

QY365AATAATTCAACAAGAAACAATACAAAAAGTGGTAGAATTACCTATCACATTTCCCAAT424

Db|||||

QY3205CTTGACTATTGAGATGCTGTTTATTATTAGTGATGAGGATTAGCACTTGATTGAAGATTCT3264

Db|||||

QY425CTTGACTATTGAGATGCTGTTTATTATTAGTGATGAGGATTAGCACTTGATTGAAGATTCT484

Db|||||

QY3265TTTAAATACTATCAGTTAAACATTTAAATATGATTATGATTAAATGTTTACTCTG3324

Db|||||

QY485TTTAAATACTATCAGTTAAACATTTAAATATGATTATGATTAAATGTTTACTCTG544

Db|||||

QY3325AGAACTGACATAAGAATCAATAAAATGATTGTTTACTCTG3365

Db|||||

QY545AGAACTGACATAAGAATCAATAAAATGATTGTTTACTCTG585

Db|||||

Search completed: January 27, 2005, 21:41:11

Job time : 1705 secs

RESULT 15

US-10-228-897-25

Sequence 25, Application US/10228897

Publication No. US20030092043A1

GENERAL INFORMATION:

APPLICANT: Fisher, Paul B.

APPLICANT: Kang, Dong-Chul

APPLICANT: Gopal Krishnan, Rahul V.

TITLE OF INVENTION: MELANOMA DIFFERENTIATION ASSOCIATED

TITLE OF INVENTION: GENE-5 AND PROMOTER AND USES THEREOF

FILE REFERENCE: A34614-A-PCT-USA (070050.2121)

CURRENT APPLICATION NUMBER: US/10/228,897

CURRENT FILING DATE: 2002-08-26

PRIOR APPLICATION NUMBER: PCT/US01/06960

PRIOR FILING DATE: 2001-02-28

PRIOR APPLICATION NUMBER: 09/515,363

PRIOR FILING DATE: 2000-02-29

NUMBER OF SEQ ID NOS: 25

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 25

LENGTH: 6406

TYPE: DNA

ORGANISM: homo sapiens

US-10-228-897-25

Query Match16.5%; Score 556; DB 14; Length 6406;

Best Local Similarity100.0%; Pred. No. 5.3e-119;

Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY1GCGGCGCGCCTGAGAGCCCTGTGGACAACTCGTCATTGTGAGCACAGAGCGGTAGAC60

Db|||||

QY5851GCGGCGCGCCTGAGAGCCCTGTGGACAACTCGTCATTGTGAGCACAGAGCGGTAGAC5910

Db|||||

QY61CCTGCTTCTCTAAGTGGGCAGCGGACAGCGCACGCACATTTCACTGTCCCAGACAA120

Db|||||

QY5911CCTGCTTCTCTAAGTGGGCAGCGGACAGCGCACGCACATTTCACTGTCCCAGACAA5970

Db|||||

QY121CAGCACCATCTGCTTGGGAGAACCCCTCTCCCTTCTCTGAGAAAGAAAGATGTCGAATGG180

Db|||||

QY5971CAGCACCATCTGCTTGGGAGAACCCCTCTCCCTTCTCTGAGAAAGAAAGATGTCGAATGG6030

Db|||||

QY181TATTCACAGACGAGAAATTTCCGCTATCTCATCTCGTGCTTCAGGCCAGGGTGAATG240

Db|||||

QY6031TATTCACAGACGAGAAATTTCCGCTATCTCATCTCGTGCTTCAGGCCAGGGTGAATG6090

Db|||||

QY241TACATCCAGGTGGAGCCTGTCTGGACTACCTGACCTTTCTGCCTGCAGAGGTGAAGAG300

Db|||||

QY6091TACATCCAGGTGGAGCCTGTCTGGACTACCTGACCTTTCTGCCTGCAGAGGTGAAGAG6150

Db|||||

QY301CAGATTCAGAGGACAGTCGCCACCTCCGGGAACATGCAGGCAGTTGAACTGCTGTGAGC360

Db|||||

QY6151CAGATTCAGAGGACAGTCGCCACCTCCGGGAACATGCAGGCAGTTGAACTGCTGTGAGC6210

Db|||||

QY361ACCTTGGAGAAGGAGTCTGGCACCTTGGTTGGACTCGGGAATTCGTGGAGGCCCTCCGG420

Db|||||

This Page Blank (uspro)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 27, 2005, 10:49:31 ; Search time 9207 Seconds
(without alignments)
13318.095 Million cell updates/sec

Title: US-09-515-363C-1
Perfect score: 3365
Sequence: 1 gcgcgcgcgcctgagagccc.....aaaatgattgtttactctg 3365

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0
Searched: 32822875 seqs, 18219865908 residues
Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:
1: gb_est1:
2: gb_est2:
3: gb_hic:
4: gb_est3:
5: gb_est4:
6: gb_est5:
7: gb_est6:
8: gb_ges1:
9: gb_ges2:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1393	41.4	2304	3 AK037057	AK037057 Mus muscu
2	1367.4	40.6	2331	3 AK040519	AK040519 Mus muscu
3	926.4	27.5	1013	5 BQ233683	BQ233683 AGENCOURT
4	900.4	26.8	1174	5 BU902097	BU902097 AGENCOURT
5	806.8	24.0	1115	4 BM467983	BM467983 AGENCOURT
6	801.8	23.8	870	5 BQ960157	BQ960157 AGENCOURT
7	728.2	21.6	918	5 BU189982	BU189982 AGENCOURT
8	718.8	21.4	755	5 BQ772836	BQ772836 UI-H-FE0-
9	711.6	21.1	1035	2 BF337464	BF337464 602035195
10	687	20.4	729	4 BF983236	BF983236 602305873
11	654.4	19.4	701	6 CA423888	CA423888 UI-H-FE1-
12	652.8	19.4	827	4 BG741146	BG741146 602631817
13	651.8	19.4	781	2 BF686405	BF686405 602143786
14	626	18.6	1046	4 BM476961	BM476961 AGENCOURT
15	625.4	18.6	627	5 BX492926	BX492926 DKF2p781C
16	622.8	18.5	1041	2 BE882040	BE882040 601505326
17	618.4	18.4	1197	3 AK018602	AK018602 Mus muscu
18	616.2	18.3	1239	4 BM467774	BM467774 AGENCOURT
19	606.4	18.0	972	6 BY720783	BY720783 BY720783
20	599.2	17.8	672	5 BQ316075	BQ316075 CM3-CT027
21	599.2	17.8	672	5 BQ316108	BQ316108 CM3-CT027
22	599.2	17.8	672	5 BQ316120	BQ316120 CM3-CT027
23	596.8	17.7	619	5 BQ016388	BQ016388 UI-H-DT1-
24	584	17.4	1611	4 BM455238	BM455238 AGENCOURT

C 25	551.2	16.4	568	5 BQ311714	BQ311714 QV3-BN004
C 26	549.8	16.3	579	5 BU685900	BU685900 UI-CF-DU1
C 27	547.8	16.3	671	6 CB453859	CB453859 709752 MA
C 28	531.4	15.8	536	1 AI718277	AI718277 a851e06.x
C 29	525.4	15.6	647	6 CB424267	CB424267 598511 MA
C 30	523.6	15.6	698	7 CO748417	CO748417 SNESTbaa6
C 31	522.4	15.5	1201	4 BM551088	BM551088 AGENCOURT
C 32	521	15.5	682	7 CF363618	CF363618 833509 MA
C 33	520.4	15.5	537	1 AI806204	AI806204 wf26c12.x
C 34	517.8	15.4	800	7 CO396840	CO396840 AGENCOURT
C 35	517	15.4	680	7 CF365022	CF365022 835621 MA
C 36	515.4	15.3	521	1 AI922705	AI922705 w011f08.x
C 37	511.2	15.2	552	1 AA134958	AA134958 2023f04.8
C 38	506.6	15.1	672	6 CA339255	CA339255 NISC 1x12
C 39	505.6	15.0	546	2 BE940626	BE940626 RC3-UT006
C 40	501.2	14.9	538	2 AW189584	AW189584 xl08h12.x
C 41	494.8	14.7	617	6 CB452989	CB452989 707909 MA
C 42	480.4	14.3	484	2 AW152541	AW152541 xf76b02.x
C 43	473	14.1	475	2 AW589567	AW589567 xo12b12.x
C 44	473	14.1	911	4 BI454996	BI454996 603173580
C 45	469.4	13.9	479	2 BF436606	BF436606 7p08d10.x

ALIGNMENTS

RESULT 1	AK037057	AK037057	2304 bp	mRNA	linear	HTC 03-APR-2004
LOCUS	AK037057	Mus musculus adult female vagina cDNA, RIKEN full-length enriched library, clone:9930105B04 product:similar to MELANOMA DIFFERENTIATION ASSOCIATED PROTEIN-5 [Homo sapiens], full insert sequence.				
DEFINITION	AK037057	HTC; CAP trapper.				
ACCESSION	AK037057	Mus musculus (house mouse)				
VERSION	AK037057.1	GI:26331913				
KEYWORDS	HTC; CAP trapper.					
SOURCE	Mus musculus					
ORGANISM	Mus musculus					
REFERENCE	1	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
AUTHORS	Carninci,P. and Hayashizaki,Y.					
TITLE	High-efficiency full-length cDNA cloning					
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)					
MEDLINE	99279253					
PUBMED	10349636					
REFERENCE	2	Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.				
AUTHORS	Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.					
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes					
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)					
MEDLINE	20499374					
PUBMED	11042159					
REFERENCE	3	Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsuai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwaki,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.				
AUTHORS	Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsuai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwaki,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.					
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer					
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)					
MEDLINE	20530913					
PUBMED	11076861					
REFERENCE	4	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.				
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.					
TITLE	Functional annotation of a full-length mouse cDNA collection					
JOURNAL	Nature 409, 685-690 (2001)					
REFERENCE	5					

AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL Nature 420, 563-573 (2002)

REFERENCE 6 (bases 1 to 2304)

AUTHORS Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.

TITLE Direct Submission

JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)

COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL:http://genome.gsc.riken.jp/

URL:http://fantom.gsc.riken.jp/.

FEATURES

source

Location/Qualifiers

1..2304

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="FANTOM,DB:9930105B04"

/db_xref="taxon:10090"

/clone="9930105B04"

/sex="female"

/tissue type="vagina"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

/dev_stage="adult"

224..>2302

CDS

/note="unnamed protein product; putative similar to MELANOMA DIFFERENTIATION ASSOCIATED PROTEIN-5 [Homo sapiens] (SPTR|Q9H3G6, evidence: FASTY, 75.6%ID, 67.6%length, match=2079)"

/codon_start=1

/protein_id="BAC29687.1"

/db_xref="GI:26331914"

/translation="MSIVCSAEDSFRNLILFPRRLKMYIQVEPVLDDLIFLSAETKE QILKINTCGNTSAEELLSTLEQGWPLGTQMFVEALEHSGNPLAARYVKPTLTDL PSPSSEAHDECLHLTLQPTLVDKLLINDVLDTCFEKGLLTVEDNRIRISAAGNSGN ESGVRELLRRIVQKENWFSTFLDLVLRQTGNDALFQELTGGGCPEDNTDLANSSTRDGP AANECLLPVDESLETAWNVDDILPEASCTDSVTSTESDTSLAEGSVSCFDESIGH NSNMGRDSGTMGSDSESVIQTRVSPPELQRPYQMEVAQPALDQKNIIICLPTGS GKTRVAVYITKDLHDKKQASESGKIVLVNKMVLAELQFRKEFNFLKKWYRIIGLS GDTQLKISFPEVVKSYDVIISTAQILENSLNLESDDDGVLSDFSLIIDECHTN KEAVYNNIMRRYLKQLRNDLKKQKPAIPLQILGLTASPGVGAKKQSEAEKHIL NTCANLDAFTIKTVKENLGQLKHQIEKPCFKFVIADDTRENPFKEKLLEIMASIQTYC QKSPMSDFGTQHYEQWAIQMEKKAADGNRKRDRVCAEHLRKYNEALQINDTIRMIDAY SHLETFTYDEKEKFAVLNDSKSDDEASSCNDQLKGDVKKSLKLDETDEFLMNLFFD NKXMLKKLAEN"

ORIGIN

Query Match 41.4%; Score 1393; DB 3; Length 2304;

Best local Similarity 79.3%; Pred. No. 0;

Matches 1691; Conservative 0; Mismatches 435; Indels 7; Gaps 3;

QY 121 CAGCACCATCTGCTTGGGAGAACCCCTCTCCC-TTCTCTGAGAAAGAAAGATGTGGAATGG 179

DB 175 CCGCCCCCGCCCGGGAGACTCCTCTCCCATTTCTTGAGACCTCAGCATGTCGATTGT 234

QY 180 GTATTCCACAGACAGAGAAATTTCCGCTATCTCATCTCTGTCGTCAGGGCCAGGGTGAAGAT 239

DB 235 CTGTTCTGCAGAGACAGCTTCAGGAATCTCATCTTATTCTTCAGGCCAGGCTGAAAAT 294

QY 240 GTACATCCAGGTGAGCCTGTGCTGGACTACCTTGACCTTTCTGCCCTGCAGAGGTGAAGGA 299

DB 295 GTACATTTCAGGTGAGCCAGTGTGGACCACCTCATCTTTCTGTCTGCAGAAACCAAGA 354

QY 300 GCAGATTTCAGAGGACAGTCGCCACCTCCGGGAACATGCAGGCAAGTTGAACCTGCTGCTGAG 359

DB 355 GCAGATTCTTAAAAAAGATCAACACCTGTGTTAAACACAGCGCGGAGAACTGCTGCTGAG 414

QY 360 CACCTTGGAGAGGGAGTCTGGCACCTTGGTTGGACTCGGGAATTCGTGGAGGCCCTCCG 419

DB 415 CACCTTGGAGCAGGACAATGGCCTCTGGATGGACGCAGATGTTCTGTGGAGGCCCTAGA 474

QY 420 GAGAACCGGCAGCCCTCTGGCCGCCGCTACATGAACCCCTGAGCTCACGGACTTGCCTC 479

DB 475 GCACAGTGGCAATCCCTAGCCGCGCTATGTCAAACCCACACTCACTGATCTGCCCTC 534

QY 480 TCCATCGTTTGAACGCTCATGATGAATATCTCCAACTGCTGAACCTCCTTCAGCCCCAC 539

DB 535 TCCTTCTCTGAGACTGCCATGACGAGTGTCTCCACTTGTGCTGACCTCTCCAGCCCCAC 594

QY 540 TCTGGTGGACAAGCTTCTAGTTAGAGACGCTCTTGATTAAGTGCATGGAGGAGGAACGTGT 599

DB 595 TTTGGTGGACAACTTCTGATTAAAGATGTTCTGGACACTTGTCTTCGAGAAGGACTATT 654

QY 600 GACAATTGAAGACAGAAACCGGATTGCTGTCGAGAAACAAATGGAATGAATCAGGTGT 659

DB 655 GACAGTCGAAGACAGAAATCGGATTTCTGTCAGGAAACAGCGGGAATGAGTCAGGTGT 714

QY 660 AAGAGAGCTACTAAAGAGATTGTGCAGAAAGAAAACTGGTTCTCTGCAATTTCTGAATGT 719

DB 715 AAGAGAGCTGCTGAGAAGGATTGTGCAGAAAGAAAACTGGTTTCTTACCTTCTCTGGATGT 774

QY 720 TCTTCGTCAAAACAGGAAACAATGAACCTTGTCAGAGATTAACAGGCTCTGATTGCTCAGA 779

DB 775 TCTGCGCCAAACTGGAAATGATGCACCTATTCCAAGAACTAACAGGTGGAGGCTGCCCAGA 834

QY 780 AAGCAATGCAGAGATTGAGAATTTATCAACAAGTTGATGGTCTCAAGTGAAGAGCAACT 839

DB 835 AGACAACACAGACTTGGCTAACTCGTCTCAGAGATGGGCTGCAGCTAATGAGTGTCT 894

QY 840 TCTTTCAACACAGTTTCAGCCAAATCTGGAGAGGAGGTCTGGGGCATGGAGATAACTC 899

DB 895 TCTGCCTGCTGTGATGAGTCAAGTCTGGAGACAGAGGCCCTGGAACTGACGACATATT 954

QY 900 ATCAGAAATCATCTTTTGCAGATTCTTCTGTAGTTTTCAGAAATCAGACACAAGTTTGGCAGA 959

DB 955 ACCAGAGGCTTCTTGTACAGATTCTGTTGACCAACAGAAATCAGACACAAGTTTGGCAGA 1014

QY 960 AGGAAGTGTGAGTGTGATGAAAGTCTTGGACATAAACAGCAACATGGGCAGTGATTC 1019

DB 1015 AGGAAGTGTGAGTGTGATGAAAGTCTTGGACATAAACAGCAACATGGGCAGGATTC 1074

QY 1020 AGGCACCATGGGAAGTGAATCAGATGAA---GAGAATGTGGCAGCAAGAGCATCCCCGA 1076

DB 1075 AGGCACCATGGGAAGTGAATCAGATGAAAGTGTATCCAGACAAAAGAGATATCCCCCGA 1134

QY 1077 GCCAGAACTCCAGCTCAGGCCCTTACCAATGGAAGTTGCCAGCCAGCCTTGAAGGGAA 1136

DB 1135 GCCAGAACTCCAGCTCAGGCCCTTACCAATGGAAGTTGCCAGCCAGCCTTAGATGGAA 1194

QY 1137 GAATATCATCATCTGCTCCCTACAGGGAGTGGAAAAACCAGAGTGGCTGTTTACATTGC 1196

DB 1195 GAATATTATTATCTGCTCTCCACAGGGAGTGGAAAAACCAGAGTGGCTGTTTACATCAC 1254

QY 1197 CAAGGATCACTTAGACAAGAAAAAGCATCTGAGCCTGGAAAAAGTTATAGTTCTTGT 1256

Db 1255 CAAAGATCACTTAGACAAGAAAGACGAGCATCTGAATCCGGGAAGGTTATCGTTCCTGT 1314

Qy 1257 CAATAAGGTACTGTAGTTGAACAGCTCTCCGCAAGGAGTTCCAACCATTTTTTGAAGAA 1316

Db 1315 CAATAAGGTAATGTAGCAGAACAACTTTCCGAAAAGAGTTCACCCCATATTTTGAAGAA 1374

Qy 1317 ATGGTATCGTGTATTGGATTAAAGTGGTGATACCCAACTGAAAATATCATTTCCAGAAGT 1376

Db 1375 ATGGTATCGAATTATTGGATTAAAGTGGCGATACCCAGCTGAAAATATCATTTCCAGAAGT 1434

Qy 1377 TGTCAAGTCCTGTGATATTATTATCAGTACAGCTCAAATCCTTGAAAACTCCCTCTTAAA 1436

Db 1435 TGTCAAATCTTACGATGTTATTATCAGCACTGCTCAAATCCTTGAAAACTCCCTCTTAAA 1494

Qy 1437 CTTGGAAAATGGAGAAGATGCTGGTGTCAATTGTGAGACTTTTCCCTCATTTATCATTGA 1496

Db 1495 TCTGGAGAGTGGAGACGATGACGGTGTGAGCTGTGAGACTTCTCTCATTTATCATTGA 1554

Qy 1497 TGAATGTCTATCACACCAACAAAGAACGAGCTGTATAATAACATCATGAGGCATTATTGAT 1556

Db 1555 TGAGTGCCATCACACCAACAAAGGAGGCGAGTCTATAACAACATCATGAGACGATATTGAA 1614

Qy 1557 GCAGAAAGTTGAAAAACAATAGACTCAAGAAAGAAAAACAACCACTGATTCCCTCTCTCA 1616

Db 1615 GCAGAAAGCTGAGAAACAATGACCTCAAGAAACAAACAAACCAACCCATTTCCCTGCCGA 1674

Qy 1617 GATACTGGGACTAACAGCTTCACCTGGTGTGGAGGGGCCACGAAGCAAGCCAAAGCTGA 1676

Db 1675 GATACTAGGACTGACAGCTTCACCTGGTGTGGAGCGAGCCAAAAGCAGTCTGAGGCTGA 1734

Qy 1677 AGAACACATTTTAAACACTATGTGCCAATCTTGTATGCTATTTACTATTAAAACTGTTAAGA 1736

Db 1735 AAAACATATTTTAAATATATGTGCCAATCTTGTATGCTCTTACCATTAAAAACAGTGAAGA 1794

Qy 1737 AAACCTTGATCAACTGAAAAACCAATACAGGAGCCCATGCAAGAAAGTTTGCCATTGCAGA 1796

Db 1795 GAATCTTGGTCAACTCAACACCAACAAATAAAGGAACCATGCAAGAAATTTGTGATTGCTGA 1854

Qy 1797 TGCAACCAGAGAGATCCATTTAAAGAAAACTTCTAGAAATTAATGACAAGGATTCAAAC 1856

Db 1855 TGACACCAGAGAAAAATCCATTTAAAGAGAAAACTTCTAGAAATTAATGCAAGCATTCAGAC 1914

Qy 1857 TTATTGTCAANTGAGTCCAATGTCTCAGATTTTGGAACTCAACCCCTATGAACAATGGGCCAT 1916

Db 1915 TTACTGCCAAAAAGTCCAATGTCTCAGATTTTGGAAACCCCAACATTATGAGCAGTGGGCCAT 1974

Qy 1917 TCAATGGAAAAAAGCTGCAAAAAAGGAAATCGCAAGAACGTTGTTGTGCAGAACA 1976

Db 1975 TCAATGGAGAAAAAAGCTGTCTAAAGACGGAATCGCAAGATCGCGTCTGTGCAGAACA 2034

Qy 1977 TTTGAGGAAGTACAATGAGGCCCTACAAATTAATGACACAAATTCGAATGATAGATGGTA 2036

Db 2035 TTTGAGGAAGTACAACGAAGCCCTACAAATCAACGACACGATCCGAATGATTGATGCATA 2094

Qy 2037 TACTCATCTTGAAACTTTCTATAATGAAGAGAAAGATAAGAAGTTTGCAGTCATAGAAGA 2096

Db 2095 TAGCCACCTGGAGACATTTCTACACTGATGAGAAAGAAAGAGTTGCGAGTCCT---CAA 2151

Qy 2097 TGATAGTGATGAGGGTGGTGATGATGATGATTTGTGATGGTGTGATGAAGATGAGGATGATT 2156

Db 2152 TGACAGCGACAGAGTGATGACGAGGCCAGCAGTTGCAATGACCAACTTAAGGGCGATGT 2211

Qy 2157 AAAGAAACCTTTGAAACTGGATGAAACAGATAGATTTCTCATGACTTTATTTTTTGA AAA 2216

Db 2212 AAAGAAATCTTTGAAACTGGACGAACCGGATGAATTTCTCATGAATTTGTTCTTGATAA 2271

Qy 2217 CAATAAAATGTTGAAAAGGCTGGCTGAAAACCC 2249

Db 2272 CAAGAAAAATGTTGAAAAAACTAGCTGAAAACCC 2304

AK040519 2331 bp mRNA linear HTC 03-APR-2004

LOCUS

DEFINITION

Mus musculus 0 day neonate thymus cDNA, RIKEN full-length enriched library, clone:A430105A06 product:similar to MELANOMA DIFFERENTIATION ASSOCIATED PROTEIN-5 [Homo sapiens], full insert sequence.

AK040519

AK040519.1 GI:26333792

HTC; CAP trapper.

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1

Carninci, P. and Hayashizaki, Y.

High-efficiency full-length cDNA cloning

Meth. Enzymol. 303, 19-44 (1999)

99279253

10349636

2

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes

Genome Res. 10 (10), 1617-1630 (2000)

20499374

11042159

3

Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer

Genome Res. 10 (11), 1757-1771 (2000)

20530913

11076861

4

The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.

Functional annotation of a full-length mouse cDNA collection

Nature 409, 685-690 (2001)

5

The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

6 (bases 1 to 2331)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

Direct Submission

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)

CDNA library was prepared and sequenced in Mouse Genome

COMMENT

Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: <http://genome.gsc.riken.jp/>

URL: <http://fantom.gsc.riken.jp/>

FEATURES

Location/Qualifiers
1..2331
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:A430105A06"
/db_xref="taxon:10090"
/clone="A430105A06"
/tissue_type="thymus"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="0 day neonate"
283..>2331
/note="unnamed protein product; putative
similar to MELANOMA DIFFERENTIATION ASSOCIATED PROTEIN-5
[Homo sapiens] (SPTR|Q9H3G6, evidence: FASTV, 75.6%ID,
67.6%length, match=2079)"
/codon_start=1
/protein_id="BAC30614.1"
/db_xref="GI:26333793"
/translation="MSIVCSAEDSFRNLILFFRPRLKMYIQVEPVLDFHILFLSAETKE
QILKINTCGNTSAEALLSTLEQGWPLGWTQMFVEALEHSGNPLAARYVKPTLIDL
PSPSETAHDECLHLLTLLOPTLVKLLINDVLTDFEKGLLTVEDNRNRIISAAGNSGN
ESGVRELLRRIQKENWFSFLDLVLRQTGNDALFQELTGGCCPENTDLANSSHRDGP
AANECLLPADVDESSETEAMVDDILPEASCTDSSTVTTESDTSLAEGSVSCFDES LGH
NSNMGRDSTGMSDSDESVIQTRVSPPELQLRPYQMEVAQPALDGKNIILCLPTGS
GKTRVAVYITKHLDDKKQASESGKVIVLNKMVLAELQFRKEFNPYLKWKYRIIGLS
GDTQLKISFPEVVKSYDVIISTAQILENSLNLESGDDGVLQSLDFSLIIIDECHHTN
KEAVNNIMRRYLKOKLRNNDLKKONKPAIPLQILGLTASPGVGAACKQSEAEKHIL
NICANLDAFTIKTVKENLGQLKHQIKEPCKEFVIADLTRENPFKEKLLIEMASIQTYC
OKSPMSDFGTQHYEOWAIQMEKKAADGNRDKRVCAEHLRKYNREALQINDTIRMIDAY
SHLETFTYDEKEKKFAVLNDSKSDDEASSCNDQLKGDVKKSLKLDDETDFELMNLFFD
N"

CDS

ORIGIN
Query Match 40.6%; Score 1367.4; DB 3; Length 2331;
Best Local Similarity 79.2%; Pred. No. 0;
Matches 1663; Conservative 0; Mismatches 431; Indels 7; Gaps 3;
QY 121 CAGCACCATCTGCTGGAGAACCTCTCCC-TTCTCTGAGAAAGAAAGATGTCGAATGG 179
DB 234 CCGCCCCGCGCGCGGAGACTCCTCTCCCATTTCCCTGAGACCTCACGATGTCGATTGT 293
QY 180 GTATTCACAGACGAGAAATTTCCGCTATCTCATCTCGTGTTCAGGGCCAGGTTGAAAT 239
DB 294 CTGTTCTGCAGAGCAGCTTCAGGAATCTCATCTTATCTTCAGGCCCCAGGTTGAAAT 353
QY 240 GTACATCCAGGTGGAGCCTGTGCTGGACTACCTGACCTTTCTGCCTGCAGAGGTGAAGGA 299
DB 354 GTACATTCAGGTGGAGCCAGTGTGACCCACCTCATCTTCTGCTGCGAAGAACCAAGA 413
QY 300 GCAGATTCAGAGGACAGTCGCCACCTCCGGGAACATGCAGGCAGTTGAACTGCTGTGAG 359
DB 414 GCAGATTTCTAAAAGATCAACACCTGTGGTAAACACCGCGCGGCAGAACTGCTGTGAG 473
QY 360 CACCTTGGAGAAGGGAGTCTGGACCTTGGTGGACTCGGGAATTCGTGGAGGCCCTCCG 419
DB 474 CACCTTGGAGCAGGGACAAATGGCCCTCTGGGATGGACGCGATGTTCTGTGGAGGCCCTAGA 533
QY 420 GAGAACCGGAGCCCTCTGGCCGCGCGCTACATGAACCTGAGCTCACGGACTTGCCTC 479
DB 534 GCACAGTGGCAATCCCTTAGCCGCGGCTATGTCAAACCCACACTCACTGATCTGCCCTC 593
QY 480 TCCATCGTTTGAGAACGCTCATGATGAATATCTCCAACCTGCTGAACCTCCTTCAGCCAC 539
DB 594 TCCTTCCTCTGAGACTGCCCATGACGAGTGTCTCCACTTGTGACCCCTCCTCCAGCCAC 653

Db 1734 GATACTAGGACTGACAGCTTCACCTGGTGTGGAGCAGCCAAAAAGCAGTCTGAGGCTGA 1793
QY 1677 AGAACACATTTTAAAACTATGTGCCAATCTTGATGCATTTACTATTAAAACTGTTAAAGA 1736
Db 1794 AAAACATATTTTAAATATATGTGCCAATCTTGATGCCTTTACCATTAAACAGTGAAGA 1853
QY 1737 AAACCTTGATCACTGAAAAACCAATACAGGAGCCATGCAAGAAGTTTGCCATTGCAGA 1796
Db 1854 GAATCTTGGTCAACTCAAAACACCAATTAAGGAACCATGCAAGAAATTTGTGATTGCTGA 1913
QY 1797 TGCAACCAGAGAGATCCATTAAAGAGAAACTTCTAGAAATTAATGACAAGGATTCAAAC 1856
Db 1914 TGACACCAGAGAAATCCATTAAAGAGAAACTTCTAGAAATTAATGGCAAGCATTCAGAC 1973
QY 1857 TTATTGTCAAATGAGTCCAATGTCAGATTTTGGAACTCAACCCCTATGAACAATGGGCCAT 1916
Db 1974 TTACTGCCAAAAAGTCCAATGTCAGATTTTGGAACTCAACCCCAACATTAATGAGCAGTGGGCCAT 2033
QY 1917 TCAAAATGGAAAAAAGCTGCAAAAAAGGAAATCGCAAGAACGTGTTTGTGCAGAAC 1976
Db 2034 TCAAAATGGAAAAAAGCTGCTAAAGACGGAATCGCAAGATCGCGTCTGTGCAGAAC 2093
QY 1977 TTTGAGGAAGTACAATGAGGCCCTACAAATTAATGACACAATTCGAATGATAGATGCGTA 2036
Db 2094 TTTGAGGAAGTACAACGAAGCCCTACAATCAACGACACCGATCGGAATGATTGATGCATA 2153
QY 2037 TACTCATCTTGAACCTTCTATAATGAAGAGAAAGATAAGAAGTTTGCAGTCATAGAAGA 2096
Db 2154 TAGCCACCTGGAGACATTTCTACACTGATGAGAAAGAAAGAAAGTTTCGAGTCCT---CAA 2210
QY 2097 TGATAGTGATGAGGCTGGTGATGATGAGTATTTGTGATGGTGATGAAGATGAGGATGATTT 2156
Db 2211 TGACAGCGACAAAGATGATGACGAGGCCAGCAGTTGCAATGACCAACTTAAGGGCGATGT 2270
QY 2157 AAAGAAACCTTTGAAACTGGATGAAACAGATAGATTTCTCATGACTTTTATTTTGA AAA 2216
Db 2271 AAAGAAATCTTTGAAACTGGACGAACCGATGAATTTCTCATGAATTTGTTCTTTGTATA 2330
QY 2217 C 2217
Db 2331 C 2331

RESULT 3
BQ233683
LOCUS
DEFINITION BQ233683 1013 bp mRNA linear EST 02-MAY-2002
AGENCOURT_7258428 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5786526
5', mRNA sequence.
ACCESSION BQ233683
VERSION BQ233683.1 GI:20415083
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1013)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM12875 row: g column: 07
High quality sequence stop: 721.
Location/Qualifiers
FEATURES
source 1. .1013

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5786526"
/tissue_type="leiomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_71"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2.1 kb."
ORIGIN
Query Match 27.5%; Score 926.4; DB 5; Length 1013;
Best Local Similarity 98.3%; Pred. No. 6.6e-215;
Matches 989; Conservative 0; Mismatches 11; Indels 6; Gaps 5;
QY 1752 GAAAAACCAATACAGGAGCCATGCAAGAAAGTTTGCCATTGCAGATGCAACCCAGAGAAGA 1811
Db 1 GAAAAACCAATACAGGAGCCATGCAAGAAAGTTTGCCATTGCAGATGCAACCCAGAGAAGA 60
QY 1812 TCCATTTAAAGAGAAACTTCTAGAAATTAATGACAAGGATTCAAACTTATTGTCAATGAG 1871
Db 61 TCCATTTAAAGAGAAACTTCTAGAAATTAATGACAAGGATTCAAACTTATTGTCAATGAG 120
QY 1872 TCCAATGTCAGATTTTGGAACTCAACCCCTATGAACAATGGGCCATTCAAATGGAAAAAA 1931
Db 121 TCCAATGTCAGATTTTGGAACTCAACCCCTATGAACAATGGGCCATTCAAATGGAAAAAA 180
QY 1932 AGCTGCAAAAAAGGAAATCGCAAGAACGTCGTTTGTGCGAGAACATTTGAGGAAGTACAA 1991
Db 181 AGCTGCAAAAAAGGAAATCGCAAGAACGTCGTTTGTGCGAGAACATTTGAGGAAGTACAA 240
QY 1992 TGAGGCCCTACAAATTAATGACACAATTCGAATGATAGATCGGTATCTCATCTTGAAAC 2051
Db 241 TGAGGCCCTACAAATTAATGACACAATTCGAATGATAGATCGGTATCTCATCTTGAAAC 300
QY 2052 TTTCTATAATGAAGAGAAAGATAAGAAAGTTTGCAGTCATAGAAGATGATAGTAGAGGG 2111
Db 301 TTTCTATAATGAAGAGAAAGATAAGAAAGTTTGCAGTCATAGAAGATGATAGTAGAGGG 360
QY 2112 TGGTGATGATGATATTGTGTGATGGTGATGAAGATGAGGATGATTTAAAGAAACCTTTGAA 2171
Db 361 TGGTGATGATGATATTGTGTGATGGTGATGAAGATGAGGATGATTTAAAGAAACCTTTGAA 420
QY 2172 ACTGGATGAAACAGATAGATTTCATGACTTTTATTTTGAACAATAAATGTTGAA 2231
Db 421 ACTGGATGAAACAGATAGATTTCATGACTTTTATTTTGAACAATAAATGTTGAA 480
QY 2232 AAGGCTGGCTGAAAAACCCAGAAATATGAAATGAAAGCTGACCAAAATTAAGAAATACCAT 2291
Db 481 AAGGCTGGCTGAAAAACCCAGAAATATGAAATGAAAGCTGACCAAAATTAAGAAATACCAT 540
QY 2292 AATGGAGCAATATACTAGGACTGAGGAATCAGCACGAGGAATAATCTTTACAAAAACACG 2351
Db 541 AATGGAGCAATATACTAGGACTGAGGAATCAGCACGAGGAATAATCTTTACAAAAACACG 600
QY 2352 ACAGAGTGCATATGCGCTTTCCAGTGGATTACTGAAAATGAAAAATTTTGTGAAGTAGG 2411
Db 601 ACAGAGTGCATATGCGCTTTCCAGTGGATTACTGAAAATGAAAAATTTTGTGAAGTAGG 660
QY 2412 AGTCAAGCCACCACCATCTGATTGGAGCTGGACACAGCAGTGAATTCAAACCCATGACACA 2471
Db 661 AGTCAAGCCACCACCATCTGATTGGAGCTGGACACAGCAGTGAATTCAAACCCATGACACA 720
QY 2472 GAATGAACAAAAAGAGTCATTAGTAAATTTCCGACCTCGAAAAATCAATCTGCTTATCGC 2531
Db 721 GAATGAACAAAAAGAGTCATTAGTAAATTTCCGACCTCGAAAAATCAATCTGCTTATCGC 780
QY 2532 TACCACAGTGGCAGAAAGTCTGGATATTTAAAGAAATGTAACATTTGTTATCCGTTATGG 2591
Db 781 TACCACAGTGGCAGAAAGTCTGGATATTTAAAGAAATGTAACATTTGTTATCCGTTATGG 840
QY 2592 TCTCGTCACCAANTGAATAGCCATGGTCCAGGCCCGTGGTCGAGCCAGAGCTGATGAGAG 2651

```
||||| 841 TCCTGTC-CAATGAATAGCATGGTCCAGGCCGTGGTCGAGCCAGACTGATGAGAG 899
||||| 2652 CACCTACGTCCTGGTTGCTCACAGTGGTTTCAGGAGTTATCGAACATGAGACAGTTAATGA 2711
||||| 900 CACCTACGTCCTGGTTGCTCACGTTGGTTTCAG--AGTATCGAACGTCGAGACA-TTAATGA 956
||||| 2712 TTTCCGAGAGAAGATGATGTATAAAGCTATACATTGTTCAAAAT 2757
||||| 957 TTTCCGAGA-AAGATGATGTAT-AAGCTATACATGCGGTTCAAAAT 1000

RESULT 4
BU902097 1174 bp mRNA linear EST 17-OCT-2002
LOCUS AGENCOURT_10127740 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:6502757
DEFINITION 5', mRNA sequence.
ACCESSION BU902097
VERSION BU902097.1 GI:24084010
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1174)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC

CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM14059 row: f column: 06
High quality sequence stop: 808.

FEATURES
    source
        1..1174
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:6502757"
            /tissue_type="leiomyosarcoma"
            /lab_host="DH10B (phage-resistant)"
            /clone_lib="NIH_MGC_71"
            /note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NotI;
            Site 2: Sali; Cloned unidirectionally. Primer: Oligo dT.
            Average insert size 2.1 kb."

ORIGIN
Query Match 26.8%; Score 900.4; DB 5; Length 1174;
Best Local Similarity 93.7%; Pred. No. 1.5e-208;
Matches 1049; Conservative 0; Mismatches 56; Indels 15; Gaps 10;

QY 1 GCGCGCCGCGCTGAGAGCCCTGTGGACAACTCGTCATTGTGAGGACAGAGCGGTAGAC 60
|||||
DB 55 GCGCGCCGCGCTGAGAGCCCTGTGGACAACTCGTCATTGTGAGGACAGAGCGGTAGAC 114
|||||

QY 61 CCTGCTTCTTAAGTGGGACGCGGACAGCGGCACGCACATTTCACCTGTCCCGCAGACAA 120
|||||
DB 115 CCTGCTTCTTAAGTGGGACGCGGACAGCGGCACGCACATTTCACCTGTCCCGCAGACAA 174
|||||

QY 121 CAGCACCATCTGCTGGGAGAACCTCTCCCTTCTCTGAGAAAGAAAGATGTCGAATGGG 180
|||||
DB 175 CAGCACCATCTGCTGGGAGAACCTCTCCCTTCTCTGAGAAAGAAAGATGTCGAATGGG 234
|||||

QY 181 TATTCACAGACGAGAAATTTCCGCTATCTCATCTCGTTCAGGGCCAGGGTGAAAATG 240
|||||
DB 235 TATTCACAGACGAGAAATTTCCGCTATCTCATCTCGTTCAGGGCCAGGGTGAAAATG 294
|||||
```

```
QY 241 TACATCCAGGTGAGCCCTGTGCTGGACTACCTGACCTTTCTGCCTGCAGAGGTGAAGGAG 300
|||||
DB 295 TACATCCAGGTGAGCCCTGTGCTGGACTACCTGACCTTTCTGCCTGCAGAGGTGAAGGAG 354
|||||

QY 301 CAGATTTCAGAGACAGTCGCCACCTCCGGGAACATGCAGGCAAGTTGAACTGCTGCTGAGC 360
|||||
DB 355 CAGATTTCAGAGACAGTCGCCACCTCCGGGAACATGCAGGCAAGTTGAACTGCTGCTGAGC 414
|||||

QY 361 ACCTTGGAGAGGAGTCTGGCACCTTGGTTGGACTCGGGAATTCGTGGAGGCCCTCCGG 420
|||||
DB 415 ACCTTGGAGAGGAGTCTGGCACCTTGGTTGGACTCGGGAATTCGTGGAGGCCCTCCGG 474
|||||

QY 421 AGAACCGGCAGCCCTCTGGCCGCCCTACATGAACCTTGAGTCAAGTCAAGGACTTGCCCTCT 480
|||||
DB 475 AGAACCGGCAGCCCTCTGGCCGCCCTACATGAACCTTGAGTCAAGTCAAGGACTTGCCCTCT 534
|||||

QY 481 CCATCGTTTGAGAACGCTCATGATGAATATCTCCAACCTGCTGAACCTTCCAGCCCACT 540
|||||
DB 535 CCATCGTTTGAGAACGCTCATGATGAATATCTCCAACCTGCTGAACCTTCCAGCCCACT 594
|||||

QY 541 CTGGTGGACAAAGCTTCTAGTTAGAGACGCTTGGATAAGTGCATGGAGGAGAACTGTTG 600
|||||
DB 595 CTGGTGGACAAAGCTTCTAGTTAGAGACGCTTGGATAAGTGCATGGAGGAGAACTGTTG 654
|||||

QY 601 ACAATTGAAGACAGAAACCGGATTGCTGTCAGAAAAACAATGAAATGAATCAGGTGTA 660
|||||
DB 655 ACAATTGAAGACAGAAACCGGATTGCTGTCAGAAAAACAATGAAATGAATCAGGTGTA 714
|||||

QY 661 AGAGAGCTACTAAAAAGGATTGTCAGAAAAAAACTGGTTCTCTGCATTCTGAATGTT 720
|||||
DB 715 AGAGAGCTACTAAAAAGGATTGTCAGAAAAAAACTGGTTCTCTGCATTCTGAATGTT 774
|||||

QY 721 CTTTCGTCAGGAGAAACAATGAACCTTGTCAGAGGTTAACAGGCTCTGATTGCTCAGAA 780
|||||
DB 775 CTTTCGTCAGGAGAAACAATGAACCTTGTCAGAGGTTAACAGGCTCTGATTGCTCAGAA 834
|||||

QY 781 AGCAATGCAGAGATTGAGAAATTTATCACAAGTTGATGGTCTCTCAAGT-GGAAGAGCAACT 839
|||||
DB 835 AGCAATGCAGAGATTGAGAAATTTATCACAAGTTGATGGTCTCTCAAGTGGGAAAGCAACT 894
|||||

QY 840 TCTTTCAACCACA-GTTTCAGCCAAATCTGG-AGAAGGAGGTTCTGGGCGATGG-AGAATAA 896
|||||
DB 895 TCTTTCAACCACAAGGTTTCAGCCAAATCTGGAAAAAGGAGGTTCTGGGCGATGGAAAAATAA 954
|||||

QY 897 CTCATC-AGAATCATCTTTTCAGATTCTTCTGTAGTTTCAGAAATCAGACACAAAGTTTG 955
|||||
DB 955 CTCATCAAAAATCATCTTTTCAGATCTCTTCTGAAGTTCCAGAACCAAAAACAAGTTTGG 1014
|||||

QY 956 C-AGAAGGAAGTGTACGCTGCTT-AGATGAAAGTCTT-GGACATAACAGCAACATGGGCA 1012
|||||
DB 1015 CAAAAAGGAAATGGCAACTGCTTAAATGAAAGTCTTGGGACATACAAACATGGGCA 1074
|||||

QY 1013 GTGATTTCAGGCACCATGGG--AAGTGATTCAGATGAAGAGAATGTGGCAGCAAGACA- 1068
|||||
DB 1075 GGGATTTCGGGCCCGGGGCAAGGATTTCAAAATAAAAAAAGGGGCCACCAAGAAA 1134
|||||

QY 1069 ---TCCCGGAGCCAGAACTCCAGCTCAGGCTTACCAA 1105
|||||
DB 1135 ATTCCCGGGAACAAAAAATTCAGCTTGGGCCCTAACAAA 1174
|||||

RESULT 5
BM467983
LOCUS AGENCOURT_6437921 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5532884
DEFINITION 5', mRNA sequence.
ACCESSION BM467983
VERSION BM467983.1 GI:18517025
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```


REFERENCE 1 (bases 1 to 1115)
NIH-MGC http://mgc.nci.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM1216 row: n column: 21
High quality sequence stop: 690.

FEATURES
source
Location/Qualifiers
1..1115
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5532884"
/tissue_type="leiomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_71"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2.1 kb."

ORIGIN
Query Match 24.0%; Score 806.8; DB 4; Length 1115;
Best Local Similarity 99.6%; Pred. No. 1.1e-185;
Matches 808; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2555 TGGATATTAAAGATGTAACATTGTTATCCGTTATGGTCTCGTCACCAATGAAATAGCCA 2614
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2 TGGATATTAAAGATGTAACATTGTTATCCGTTATGGTCTCGTCACCAATGAAATAGCCA 61
QY 2615 TGGTCCAGGCCCGTGGTCGAGCCGAGCTGATGAGAGCACCTACGTCCTGGTTGCTCACA 2674
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
62 TGGTCCAGGCCCGTGGTCGAGCCGAGCTGATGAGAGCACCTACGTCCTGGTTGCTCACA 121
QY 2675 GTGGTTCAGGAGTTATCGAACATGAGACAGTTAATGATTTCCGAGAGAAATGATGTATA 2734
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
122 GTGGTTCAGGAGTTATCGAACGTGAGACAGTTAATGATTTCCGAGAGAAATGATGTATA 181
QY 2735 AAGCTATACATTGTGTTCAAAATATGAACACGAGGAGTATGCTCATAGATTTTGGAAAT 2794
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
182 AAGCTATACATTGTGTTCAAAATATGAACACGAGGAGTATGCTCATAGATTTTGGAAAT 241
QY 2795 TACAGATGCAAACTATAATGGAAAAGAAAATGAAACCAAGAGAAAATATTGCCAAGCATT 2854
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
242 TACAGATGCAAACTATAATGGAAAAGAAAATGAAACCAAGAGAAAATATTGCCAAGCATT 301
QY 2855 ACAAGAATAACCCATCACTAATAACTTTCCCTTTGCAAAAACATGCAGTGTGCTAGCCCTGTT 2914
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
302 ACAAGAATAACCCATCACTAATAACTTTCCCTTTGCAAAAACATGCAGTGTGCTAGCCCTGTT 361
QY 2915 CTGGGGAAGATATCCATGTAATTGAGAAAATGCATCACGTCAATATGACCCCGAATTCA 2974
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
362 CTGGGGAAGATATCCATGTAATTGAGAAAATGCATCACGTCAATATGACCCCGAATTCA 421
QY 2975 AGGAACCTTTACATTGTAAGAGAAAACAAAGCACTGCAAAAAGAGTGTGCCGACTATCAAA 3034
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
422 AGGAACCTTTACATTGTAAGAGAAAACAAAGCACTGCAAAAAGAGTGTGCCGACTATCAAA 481
QY 3035 TAAATGGTGAATCATCTGCAAAATGTGGCAGGCTTGGGGAACAATGATGGTGCAACAAG 3094
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
482 TAAATGGTGAATCATCTGCAAAATGTGGCAGGCTTGGGGAACAATGATGGTGCAACAAG 541
QY 3095 GCTTAGATTTGCCCTTGCTCAAAAATAAGGAATTTTGTAGTGGTTTCAAAAATAATTCAA 3154
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
542 GCTTAGATTTGCCCTTGCTCAAAAATAAGGAATTTTGTAGTGGTTTCAAAAATAATTCAA 601

QY 3155 CAAAGAAACAATACAAAAGTGGGTAGAATTACCTATCACATTTCCCAATCTTGACTATT 3214
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
602 CAAAGAAACAATACAAAAGTGGGTAGAATTACCTATCACATTTCCCAATCTTGACTATT 661
QY 3215 CAGAATGCTGTTTATTATTAGTATGAGGATTAGCACCTTGATTGAAGATTCTTTTAAATAC 3274
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
662 CAGAATGCTGTTTATTATTAGTATGAGGATTAGCACCTTGATTGAAGATTCTTTTAAATAC 721
QY 3275 TATCAGTTAAACATTTAATATGATTATGATTAATGATTATCATGCTACAGAACTGACA 3334
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
722 TATCAGTTNAACATTTAATATGATTATGATTAATGATTATCATGCTACAGAACTGACA 781
QY 3335 TAAGAATCAATAAAATGATTGTTTACTCTG 3365
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
782 TAAGAATCAATAAAATGATTGTTTACTCTG 812

RESULT 6
BQ960157
LOCUS BQ960157 870 bp mRNA linear EST 21-AUG-2002
DEFINITION AGENCOURT 8949486 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:6472264
5', mRNA sequence.
ACCESSION BQ960157
VERSION BQ960157.1 GI:22375635
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 870)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM14005 row: o column: 17
High quality sequence stop: 738.

FEATURES
source
Location/Qualifiers
1..870
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6472264"
/tissue_type="leiomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_71"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2.1 kb."

ORIGIN
Query Match 23.8%; Score 801.8; DB 5; Length 870;
Best Local Similarity 98.4%; Pred. No. 1.8e-184;
Matches 851; Conservative 0; Mismatches 9; Indels 5; Gaps 4;

QY 1940 AAAAAGGAAATCGCAAAGAACGTTTGTGCAGAACATTTGAGGAAGTACAATGAGGCC 1999
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1 AAGAAGGAAATCGCAAAGAACGTTTGTGCAGAA-ATTGAGGAAGTACAATGAGGCC 59
QY 2000 TACAAATTAAATGACACAATTCGAATGATAGATGCGTATACTCATCTTGAAACTTTCTATA 2059
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
60 TACAAATTAAATGACACAATTCGAATGATAGATGCGTATACTCATCTTGAAACTTTCTATA 119
QY 2060 ATGAAGAGAAAGATAAGAAGTTTGCAGTCATAGAAGATGATGATGAGGGTGGTGATG 2119
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
120 ATGAAGAGAAAGATAAGAAGTTTGCAGTCATAGAAGATGATGATGAGGGTGGTGATG 179

found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13546 row: a column: 17
High quality sequence stop: 556.
Location/Qualifiers

FEATURES

source
1. .918
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6173848"
/tissue_type="leiomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC_71"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: Sali; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2.1 kb. "

ORIGIN

Query Match	21.6%;	Score 728.2;	DB 5;	Length 918;
Best Local Similarity	97.5%;	Pred. No. 1.7e-166;		
Matches 814;	Conservative 0;	Mismatches 13;	Indels 8;	Gaps 7;
Qy	1245	TATAGTTCTTGTCAATAAGGTACTGCTAGTTGAACAGCTCTTCCGCAAGGAGTTCACACC	1304	
Db	1	TATAGTTCTTGTCAATAAGGTACTGCTAGTTGAACAGCTCTTCCGCAAGGAGTTCACACC	60	
Qy	1305	ATTTTGAAGAAATGGTATCGTGTATTGGATTAAAGTGGTGATACCCAACTGAAATATC	1364	
Db	61	ATTTTGAAGAAATGGTATCGTGTATTGGATTAAAGTGGTGATACCCAACTGAAATATC	120	
Qy	1365	ATTTCCAGAAATGTCAAGTCTCTGTGATATTATTATCAGTACAGCTCAAAATCCTTGAAAA	1424	
Db	121	ATTTCCAGAAATGTCAAGTCTCTGTGATATTATTATCAGTACAGCTCAAAATCCTTGAAAA	180	
Qy	1425	CTCCCTCTTAAACTTGGAAAAATGGAGAAAGATGCTGGTGTCAATTGTCTCAGACTTTTCCCT	1484	
Db	181	CTCCCTCTTAAACTTGGAAAAATGGAGAAAGATGCTGGTGTCAATTGTCTCAGACTTTTCCCT	240	
Qy	1485	CATTATCATTGATGAATGTCTATCACACCAACAAAGAGCAGTGTATATAATACATCATGAG	1544	
Db	241	CATTATCATTGATGAATGTCTATCACACCAACAAAGAGCAGTGTATATAATACATCATGAG	300	
Qy	1545	GCATTATTTGATGCAGAAAGTTGAAAAACAATAGACTCAAGAAAGAAAAACAACCCAGTGAT	1604	
Db	301	GCATTATTTGATGCAGAAAGTTGAAAAACAATAGACTCAAGAAAGAAAAACAACCCAGTGAT	360	
Qy	1605	TCCCTCTCTCAGATACTGGGACTAACAGCTTCACTGGTGTGGAGGGGCCACGAAGCA	1664	
Db	361	TCCCTCTCTCAGATACTGGGACTAACAGCTTCACTGGTGTGGAGGGGCCACGAAGCA	420	
Qy	1665	AGCCAAAGCTGAAGAACACATTTTAAACTATGTGCCAATCTTGTGATGCTTTACTATTAA	1724	
Db	421	AGCCAAAGCTGAAGAACACATTTTAAACTATGTGCCAATCTTGTGATGCTTTACTATTAA	480	
Qy	1725	AACGTGTTAAAGAAAAACCTTGATCAACTGAAAAACCAATACAGGAGCCATGCAAGAAGTT	1784	
Db	481	AACGTGTTAAAGAAAAACCTTGATCAACTGAAAAACCAATACAGGAGCCATGCAAGAAGTT	540	
Qy	1785	TGCCATTGCAGATGCAACCCAGAGAAGATCCATTAAAGAGAAAACTTCTAGAAATAATGAC	1844	
Db	541	TGCCATTGCAGATGCAACCCAGAGAAGATCCATTAAAGAGAAAACTTCTAGAAATAATGAC	600	
Qy	1845	AAGGATTCAAAACCTTATTGTCAAATGAGTCCCAATGTCTCAGATTTTGGAACTCAACCTATGA	1904	
Db	601	AAGGATTCAAAACCTTATTGTCAAATGAGTCCCAATGTCTCAGATTTTGGAACTCAACCTATGA	660	
Qy	1905	ACAATGGG-CCATTCAAATGGAAAAA-----GCTGCAAAAAAAGGAATCGCAAGAACG	1961	
Db	661	AACATGGGCCCATTCCAATGGAAAAAAGGCTGCAAAAAAGGAATCGCAAGAACG	720	
Qy	1962	TG-TTTGTGCAGAACATTTGAGGAAGTACAATG-AGGCCCTACAAATT-AATGACACAAT	2018	
Db	721	TGTTTGTGCAAAACATTTGAGGAAGTACAATGAAGGCCCTACAAATTAAATGACACAAT	780	

Qy	2120	ATGAGTATTGTGATGGTGAAGATGAGGATGATTTAAAGAAACCTTTGAAACTGGATG	2179	
Db	180	ATGAGTATTGTGATGGTGAAGATGAGGATGATTTAAAGAAACCTTTGAAACTGGATG	239	
Qy	2180	AAACAGATAGATTCTCATGACTTTATTTTGAAGAAACAATAAATGTTGAAAGGCTGG	2239	
Db	240	AAACAGATAGATTCTCATGACTTTATTTTGAAGAAACAATAAATGTTGAAAGGCTGG	299	
Qy	2240	CTGAAACCCAGAAATATGAATAATGAAGCTGACCAAAATTAAGAAATACCATATGGAGC	2299	
Db	300	CTGAAACCCAGAAATATGAATAATGAAGCTGACCAAAATTAAGAAATACCATATGGAGC	359	
Qy	2300	AATATACTAGGACTGAGGAATCAGCACGAGGAATAATCTTTACAAAAACACGACAGAGTG	2359	
Db	360	AATATACTAGGACTGAGGAATCAGCACGAGGAATAATCTTTACAAAAACACGACAGAGTG	419	
Qy	2360	CATATCGCTTCCCAGTGGATTACTGAAAAATGAAAAATTTGCTGAAGTAGGAGTCAAAG	2419	
Db	420	CATATCGCTTCCCAGTGGATTACTGAAAAATGAAAAATTTGCTGAAGTAGGAGTCAAAG	479	
Qy	2420	CCCACCATCTGATTGGAGCTGGACACAGCAGTGAGTTCAAACCCATGACACAGAATGAAC	2479	
Db	480	CCCACCATCTGATTGGAGCTGGACACAGCAGTGAGTTCAAACCCATGACACAGAATGAAC	539	
Qy	2480	AAAAGAAAGTCATTAGTAAATTTTCGCACTGGAAAAATCAATCTGCTTATCGCTACCCAG	2539	
Db	540	AAAAGAAAGTCATTAGTAAATTTTCGCACTGGAAAAATCAATCTGCTTATCGCTACCCAG	599	
Qy	2540	TGGCAGAAAGAGTCTGGATATTAAAGAAATGTAACATTTGTTATCCGTTATGGTCTCGTCA	2599	
Db	600	TGGCAGAAAGAGTCTGGATATTAAAGAAATGTAACATTTGTTATCCGTTATGGTCTCGTCA	659	
Qy	2600	CCAATGAAATAGCCATGGTCCAGGCCCGTGGTCGAGCCAGAGCTGATGAGAGCACCTACG	2659	
Db	660	CCAATGAAATAGCCATGGTCCAGGCCCGTGGTCGAGCCAGAGCTGATGAGAGCACCTACG	719	
Qy	2660	TCCTGGTTGCTCACAGT-GGTTCAGGAGTTATCGAACATGAGACAGTTAATGATTTCCGA	2718	
Db	720	TCCTGGTTGCTCACAGTGGGTTCCAGGAGTTATCGAACGTGAGACAGTTAATGATTTCCGA	779	
Qy	2719	GAGAAGATGATGTATAAAGCTATACATTGTGTTCAAAATATG-AAACCAGAGGAGTAT--	2775	
Db	780	GAGAAGATGATGTATAAAGCTATACATTGTGTTCAAAATATGAAACCCAGAGGAGTATGG	839	
Qy	2776	GCTCATAAGATTTTGGAAATTACAGA	2800	
Db	840	GCTCATAAGATTTTGGAAATTNACA	864	

RESULT 7
BU189982
LOCUS
DEFINITION BU189982 918 bp mRNA linear EST 04-SEP-2002
5' mRNA sequence.
ACCESSION BU189982
VERSION BU189982.1 GI:22703966
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 918)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be

QY	2019	TCGAAT-GATAGATGCGTATACATC-TTGAACTTTCTATAATGAGAGAAAG	2071
Db	781	TTCAATGGATAAATGCGTATACATCTTTGAACTTTCTATAATGAGGAAAG	835
RESULT 8			
BQ772836/c			
LOCUS	BQ772836 755 bp mRNA linear EST 26-JUL-2002		
DEFINITION	UI-H-FEO-bbn-c-04-0-UI.s1 NCI CGAP_FEO Homo sapiens cDNA clone		
ACCESSION	UI-H-FEO-bbn-c-04-0-UI 3', mRNA sequence.		
VERSION	BQ772836		
KEYWORDS	BQ772836.1 GI:21981312		
SOURCE	EST.		
ORGANISM	Homo sapiens (human)		
REFERENCE	Homo sapiens		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	1 (bases 1 to 755)		
JOURNAL	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap .		
COMMENT	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index		
UNPUBLISHED (1997)	Unpublished (1997)		
CONTACT: Robert Strausberg, Ph.D.	Contact: Robert Strausberg, Ph.D.		
Email: cgapbs-r@mail.nih.gov	Email: cgapbs-r@mail.nih.gov		
Tissue Procurement: James Martin	Tissue Procurement: James Martin		
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa	cDNA Library preparation: Dr. M. Bento Soares, University of Iowa		
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa	cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa		
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa	DNA Sequencing by: Dr. M. Bento Soares, University of Iowa		
Clone Distribution: Clone distribution information can be obtained from Dr. M. Bento Soares, bento-soares@uiowa.edu	Clone Distribution: Clone distribution information can be obtained from Dr. M. Bento Soares, bento-soares@uiowa.edu		
Seq primer: M13 FORWARD	Seq primer: M13 FORWARD		
POLYA=Yes.	POLYA=Yes.		
FEATURES			
source	Location/Qualifiers		
	1..755		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="UI-H-FEO-bbn-c-04-0-UI"		
	/tissue_type="Chondrosarcoma Cell line"		
	/dev_stage="Adult"		
	/lab_host="DH10B (Life Technologies)"		
	/clone_lib="NCI CGAP_FEO"		
	/note="Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site_1: EcoR I; Site_2: Not I; NCI CGAP_FEO is a cDNA library containing the following tissue(s): a pool of 3 chondrosarcoma cell lines (grade 2) The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is CGCTACGGAC. The cell lines was provided by Dr James Martin of University of Iowa.		
	TAG_TISSUE=Human grade 2 chondrosarcoma cell line pool		
	TAG_LIB=UI-H-FEO		
	TAG_SEQ=CGCTACGGAC"		
ORIGIN			
Query Match	21.4%; Score 718.8; DB 5; Length 755;		
Best Local Similarity	99.3%; Pred. No. 3.3e-164;		
Matches 731;	Conservative 0; Mismatches 4; Indels . 1; Gaps 1;		
QY	2630	GTCGAGCCAGAGCTGATGAGAGCACCTACGTCCTGCTGCTCACAGTGGTTCCAGGAGTTA	2689
Db	754	GTCGAGCCAGAGCTGATGAGAGCACCTACGTCCT-GTTGCTCACAGTGTTCAGGAGTAN	696
QY	2690	TCGAACATGAGACAGTTAATGATTTCCGAGAGAGAGATGATGTATAAAGCTATACATTGTG	2749

High quality sequence start: 5
High quality sequence stop: 695.
Location/Qualifiers
1..1035
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4183126"
/tissue_type="glioblastoma with EGFR amplification"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI CGAP Brn64"
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: Sall; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.57 kb. Constructed by Life
Technologies. Note: this is a NCI CGAP Library."

ORIGIN

Query Match 21.1%; Score 711.6; DB 2; Length 1035;
Best Local Similarity 97.2%; Pred. No. 2e-162;
Matches 756; Conservative 0; Mismatches 19; Indels 3; Gaps 3;

QY 1049 AGAATGTGGCAGCAAGAGCATCCCGGAGCCAGAACTCCAGCTCAGGCCTTACCAAATGG 1108
|||||
Db 4 AGAATGTGGCAGCAAGAGCATCCCGGAGCCAGAACTCCAGCTCAGGCCTTACCAAATGG 63

QY 1109 AAGTTGCCAGCCAGCCTTGGAGGGAAGAATATCATCATCTGCCTCCCTACAGGGAGTG 1168
|||||
Db 64 AAGTTGCCAGCCAGCCTTGGAGGGAAGAATATCATCATCTGCCTCCCTACAGGGAGTG 123

QY 1169 GAAAAACCAGAGTGGCTGTTTACATTGCCAAGGATCACTTAGACAAGAGAAAAAGCAT 1228
|||||
Db 124 GAAAAACCAGAGTGGCTGTTTACATTGCCAAGGATCACTTAGACAAGAGAAAAAGCAT 183

QY 1229 CTGAGCTGGAAAAGTTATAGTTCTTGTCAAATAAGGTACTGCTAGTTGAACAGCTCTTCC 1288
|||||
Db 184 CTGAGCTGGAAAAGTTATAGTTCTTGTCAAATAAGGTACTGCTAGTTGAACAGCTCTTCC 243

QY 1289 GCAAGGAGTTCCAAACCATTTTGAAGAAATGGTATCGTGTATTGGATTAAAGTGGTGATA 1348
|||||
Db 244 GCAAGGAGTTCCAAACCATTTTGAAGAAATGGTATCGTGTATTGGATTAAAGTGGTGATA 303

QY 1349 CCCAACTGAAAAATATCATTTCCAGAAAGTTGTCAAAGTCTGTGATATTATTATCAGTACAG 1408
|||||
Db 304 CCCAACTGAAAAATATCATTTCCAGAAAGTTGTCAAAGTCTGTGATATTATTATCAGTACAG 363

QY 1409 CTCAATCTTGAAGAACTCCCTCTTAAACTTGGAAATGGAGAAGATGCTGTTTCAAT 1468
|||||
Db 364 CTCAATCTTGAAGAACTCCCTCTTAAACTTGGAAATGGAGAAGATGCTGTTTCAAT 423

QY 1469 TGTCAGACTTTTCCCTCATTTATCATTTGATGAATGTCATCACACCAACAAAGAACGAGTGT 1528
|||||
Db 424 TGTCAGACTTTTCCCTCATTTATCATTTGATGAATGTCATCACACCAACAAAGAACGAGTGT 483

QY 1529 ATAATAACATCATGAGGCATTATTGATGCAGAAAGTTGAAAAACAATAGACTCAAGAAAG 1588
|||||
Db 484 ATAATAACATCATGAGGCATTATTGATGCAGAAAGTTGAAAAACAATAGACTCAAGAAAG 543

QY 1589 AAAACAAACCAGTGATTTCCCTTCCCTCAGATACTGGGACTAACAGCTTACCTGGTGTG 1648
|||||
Db 544 AAAACAAACCAGTGATTTCCCTTCCCTCAGATACTGGGACTAACAGCTTACCTGGTGTG 603

QY 1649 GAGGGCCACGAAGCAAGCCAAAGCTGAAGAACACATTTTAAAACTATGTGCCAATCTTG 1708
|||||
Db 604 GAGGGCCACGAAGCAAGCCAAAGCTGAAGAACACATTTTAAAACTATGTGCCAATCT-TG 662

QY 1709 ATGCAFTTACTATTAAAACTGTTAAAGAAAAACCTTGATCAACTGAAAAACCAAATACAGG 1768
|||||
Db 663 ATGCAFTTACTATTAAAACTGTTAAAGAAAAACCTTGATCACTGGGAAAAACCAAATACAGA 722

QY 1769 AGCCATGCAAGAAAGTTTGCCATTGCAGATGCAACCCAGAGAAGATCCATTTAAAGAGAA 1826
|||||
Db 723 GGCCATGCGAGAAG-TTGGCGCTGCGGATGCGACC-GAGAAGATCCCTTTAAGGGAAA 778

RESULT 10
BF983236
LOCUS
DEFINITION
602305873F1 NIH_MGC_88 Homo sapiens cDNA clone IMAGE:4397083 5',
mRNA sequence.
BF983236
VERSION
BF983236.1 GI:12386048
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
1 (bases 1 to 729)
AUTHORS
NIH-MGC http://mgi.nci.nih.gov/.
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10097 row: a column: 20
High quality sequence stop: 665.
Location/Qualifiers
1..729
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4397083"
/tissue_type="duodenal adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 88"
/note="Organ: small intestine; Vector: pCMV-SPORT6;
Site_1: NotI; Site_2: Sall; Cloned unidirectionally;
oligo-dt primed. Average insert size 1.767 kb. Library
enriched for full-length clones and constructed by Life
Technologies. Note: this is a NIH_MGC Library."

ORIGIN

Query Match 20.4%; Score 687; DB 4; Length 729;
Best Local Similarity 98.9%; Pred. No. 1.9e-156;
Matches 723; Conservative 0; Mismatches 5; Indels 3; Gaps 3;

QY 586 GAGGAGGAACTGTTGACAATTGAAGACAGAAACCGGATTGCTGTCAGAAAAACAATGGA 645
|||||
Db 1 GAGGAGGAACTGTTGACAATTGAAGACAGAAACCGGATTGCTGTCAGAAAAACAATGGA 60

QY 646 AATGAATCAGGTGTAAGAGAGCTACTAAAAAGGATTGTGCAGAAAGAAAACTGTTCTCT 705
|||||
Db 61 AATGAATCAGGTGTAAGAGAGCTACTAAAAAGGATTGTGCAGAAAGAAAACTGTTCTCT 120

QY 706 GCATTTCTGAATGTTCTCGTCAACAGGAAACAATGAACCTTGTCACAGATTGATGGTCTCAA 825
|||||
Db 121 GCATTTCTGAATGTTCTCGTCAACAGGAAACAATGAACCTTGTCACAGATTGATGGTCTCAA 180

QY 766 TCTGATTGCTCAGAAAACAATGCAGAGATTGAGAAATTTATCACAAGTTGATGGTCTCAA 825
|||||
Db 181 TCTGATTGCTCAGAAAACAATGCAGAGATTGAGAAATTTATCACAAGTTGATGGTCTCAA 240

QY 826 GTGGAAGAGCAACTTTTCAACCACAGTTTCAGCCAAATCTGGAGAGGAGGTCTGGGGC 885
|||||
Db 241 GTGGAAGAGCAACTTTTCAACCACAGTTTCAGCCAAATCTGGAGAGGAGGTCTGGGGC 300

QY 886 ATGGAGAATAACTCATCAGAATCATCTTTTCAGAGATTCTTCTGTAGTTTCAGAAATCAGAC 945
|||||
Db 301 ATGGAGAATAACTCATCAGAATCATCTTTTCAGAGATTCTTCTGTAGTTTCAGAAATCAGAC 360

QY 946 ACAAGTTTGGCAGAGGAAGTGTGCTGCTTAGATGAAAGTCTTGGACATAACAGCAAC 1005

KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
1 (bases 1 to 827)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: James Cleaver, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA
Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10630 row: b column: 06
High quality sequence stop: 763.
FEATURES
source
1..827
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4776869"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI CGAP Skn3"
/note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI CGAP Library."
ORIGIN
Query Match 19.4%; Score 652.8; DB 4; Length 827;
Best Local Similarity 96.1%; Pred. No. 4.6e-148;
Matches 798; Conservative 0; Mismatches 17; Indels 15; Gaps 12;
QY 426 CGGAGCCCTCTGGCCGCCGTACATGAATATCTCCAACCTGCTGAACCTCTCCAG-CCCACCTCTGG 544
Db 1 CGGAGCCCTCTGGCCGCCGTACATGAACCTCTGAACCTCTCCAGCTTGCCTCTCCATC 59
QY 486 GTTTGAGAACGCTCATGATGAATATCTCCAACCTGCTGAACCTCTCCAG-CCCACCTCTGG 544
Db 60 GTTTGAGAACGCTCATGATGAATATCTCCAACCTGCTGAACCTCTCTCAGTCCCACTCTGG 119
QY 545 TGGACAAGCTTCTAGTTAGAGACGCTCTTGGATAAGTGCATGGAGGAGGAAGTGTGACAA 604
Db 120 TGGACAAGCTTCTAGTTAGAGACGCTCTTGGATAAGTGCATGGAGGAGGAAGTGTGACAA 179
QY 605 TTGAAGACAGAAACCGGATTGCTGCTGCAGAAACAAATGGAATGAATCAGGTGTAAGAG 664
Db 180 TTGAAGACAGAAACCGGATTGCTGCTGCAGAAACAAATGGAATGAATCAGGTGTAAGAG 239
QY 665 AGCTACTAAAAGGATTGTGCAGAAAGAAAACCTGGTTCTCTGCATTTCTGAATGTTCTTC 724
Db 240 AGCTACTAAAAGGATTGTGCAGAAAGAAAACCTGGTTCTCTGCATTTCTGAATGTTCTTC 299
QY 725 GTCAACAGGAAACAATGAATCTGTCCAAGAGTTAACAGGCTCTGATTG--CTCAGAAAG 782
Db 300 GTCAACAGGAAACAATGAATCTGTCCAAGAGTTAACAGGCTCTGATTGGCTCAGAAAG 359
QY 783 CAATGACAGAGATTGAGAATTTATCACAAGTTGATGGTCTCTCAAGTGGAGGACCACTTCT 842
Db 360 CAATGACAGAGATTGAGAATTTATCACAAGTTGATGGTCTCTCAAGTGGAGGACCACTTCT 419
QY 843 TTCACCACAGTTTCAGCCCAATCTGGA-GAAGGAGTCTGGGGCATGGAGATAACTCAT 901
Db 420 TTCACCACAGTTTCAGCCCAATCTGGACGAAGGAGTCTGGGGCATGGAGATAACTCAT 479
QY 902 CAGAAATCATCTTTTGCA-GATTCTTCTGTAGTTTTCAGAAATCAGACACAAAGTTTGGCA--G 958
Db 480 CAGAAATCATCTTTTGCAAGATTCTTCTGTAGTTTTCAGAAATCAGACACAAAGTTTGGCACGC 539

QY 959 AAGGAAGTGTGAGCTGCTTAGATGAAAGTCTTGA-CATAACAGCAACATGGGCAGTGAT 1017
Db 540 AAGGAAGTGTGAGCTGCTTAGATGAAAGTCTTGGACCATACAGCAACATGGGCAGTGAT 599
QY 1018 TCAGGCACCATGGGAAGTGAATTCAGATGA-GAGAATGTGGCAGCAAGAGCATCCCCGGA 1076
Db 600 TCAGGCACCATGGGAAGTGAATTCAGATGAACGAGAATGTGGCAGCAAGAGCATCCCCGGA 659
QY 1077 GCCAGAACTCCAGCTCAGGCCTTACCAAAATGG--AAGTTGCCAGCCAGCCTTGAAGGG 1134
Db 660 GCCAGAACTCCAGCTCAGGCCTTACCAAAATGGCAGCATTTGCCAGCCAG-CTTGGCAGGG 718
QY 1135 AAGAATATCATCTGCTCCCTACAGGGAGTGGAAAAACCAAGAGTGGCTGTTTACATT 1194
Db 719 AAGACTATCATCTG-CTCCCTACAGGGAGTGGAAAAACCAAGAGTGGCTGATTACATT 777
QY 1195 -GCCAAGGATCATTAGACAAGAAGAAAAAAGCATCTGAGCCTGGAAAAG 1243
Db 778 AGCCACGGATCATTAGACAAGAAGAACAAAGCATCTGAGCCTGGAAACAG 827
RESULT 13
BF686405
LOCUS BF686405
DEFINITION BF686405 781 bp mRNA linear EST 22-DEC-2000
mRNA sequence.
ACCESSION BF686405
VERSION BF686405.1 GI:11971813
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 781)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1170 row: 1 column: 22
High quality sequence stop: 731.
FEATURES
source
1..781
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4304805"
/tissue_type="leiomyosarcoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_46"
/note="Organ: uterus; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."
ORIGIN
Query Match 19.4%; Score 651.8; DB 2; Length 781;
Best Local Similarity 99.4%; Pred. No. 7.9e-148;
Matches 675; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

QY 385 CTTGGTTGGACTCGGGAATTCGTGGAGGCCCTCCGGAGAACCGGCAGCCCTCTGGCGGCC 444
Db 5 CCTGGTGGACTCGGGAATTCGTGGAGGCCCTCCGGAGAACCGGCAGCCCTCTGGCGGCC 64
QY 445 CGCTACATGAACCTGAGCTCACGGACTTGCCCTCTCCATCGTTTGAGAACGCTCATGAT 504
Db 65 CGCTACATGAACCTGAGCTCACGGACTTGCCCTCTCCATCGTTTGAGAACGCTCATGAT 124
QY 505 GAATATCTCCAACTGCTGAACCTCTTCAGCCCACTCTGGTGACAAGCTTCTAGTTAGA 564
Db 125 GAATATCTCCAACTGCTGAACCTCTTCAGCCCACTCTGGTGACAAGCTTCTAGTTAGA 184
QY 565 GACGTCTTGGATAAGTGCATGGAGGAGAACTGTTGACAAATTGAAGACAGAAACCGGATT 624
Db 185 GACGTC-TGGATAAGTGCATGGAGGAGAACTGTTGACAAATTGAAGACAGAAACCGGATT 243
QY 625 GCTGCTGCAGAAAAACAATGGAAATGAATCAGGTGTAAGAGAGCTACTAAAAAGGATTGTG 684
Db 244 GCTGCTGCAGAAAAACAATGGAAATGAATCAGGTGTAAGAGAGCTACTAAAAAGGATTGTG 303
QY 685 CAGAAAGAAAACCTGTTCTCTGCATTTCTGAATGTTCTTCGTCAAAACAGGAAACAATGAA 744
Db 304 CAGAAAGAAAACCTGTTCTCTGCATTTCTGAATGTTCTTCGTCAAAACAGGAAACAATGAA 363
QY 745 CTTGTCCAAGAGTTAAACAGGCTCTGATTGCTCAGAAAGCAATGCAGAGATTGAGAAATTA 804
Db 364 CTTGTCCAAGAGTTAAACAGGCTCTGATTGCTCAGAAAGCAATGCAGAGATTGAGAAATTA 423
QY 805 TCACAAAGTTGATGCTCCTCAAGTGGAGAGCAACTCTTTCAACCACAGTTTCAGCCAAAT 864
Db 424 TCACAAAGTTGATGCTCCTCAAGTGGAGAGCAACTCTTTCAACCACAGTTTCAGCCAAAT 483
QY 865 CTGGAGAAGGAGTCTGGGCGATGGAGAAATAACTCATCAGAAATCATCTTTTGCAGATTCT 924
Db 484 CTGGAGAAGGAGTCTGGGCGATGGAGAAATAACTCATCAGAAATCATCTTTTGCAGATTCT 543
QY 925 TCTGTAGTTTCAGAAATCAGACACAAAGTTTGGCAGAAAGTGTGAGTGTGCTTATAGTAA 984
Db 544 TCTGTAGTTTCAGAAATCAGACACAAAGTTTGGCAGAAAGTGTGAGTGTGCTTATAGTAA 603
QY 985 AGTCTTGGACATAACAG-CAACATGGGCGAGTGATTTCAGGCACCATGGGAAGTGATTGATA 1043
Db 604 AGTCTTGGACATAACAGACAAACATGGGCGAGTGATTTCAGGCACCATGGGAAGTGATTGATA 663
QY 1044 TGAAGAGAAATGTGGCAGCA 1062
Db 664 TGAAGAGAAATGTGGCAGCA 682

RESULT 14
BM476961
LOCUS
DEFINITION
AGENCOURT_6481569 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:555523
5', mRNA sequence.
BM476961
VERSION
BM476961.1 GI:18526003
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1046)
NIH-MGC <http://mgi.nci.nih.gov/>
AUTHORS
National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE
Unpublished (1999)
JOURNAL
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov
Plate: LLAM12275 row: n column: 04
High quality sequence stop: 622.
FEATURES
source
location/Qualifiers
1..1046
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:555523"
/tissue_type="leiomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_71"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2.1 kb. "

ORIGIN

Query Match 18.6%; Score 626; DB 4; Length 1046;
Best Local Similarity 97.7%; Pred. No. 1.7e-141;
Matches 635; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY 172 TCGAATGGGTATTCACAGACAGAGAAATTTCCGCTATCTCATCTCGTGTTCAGGGCCAGG 231
Db 1 TCGAATGGGTATTCACAGACAGAGAAATTTCCGCTATCTCATCTCGTGTTCAGGGCCAGG 60
QY 232 GTGAAATGTACATCCAGGTGAGGCCTGTGTGGACTACCTGACCTTTCTGCGTCGAGAG 291
Db 61 GTGAAATGTACATCCAGGTGAGGCCTGTGTGGACTACCTGACCTTTCTGCGTCGAGAG 120
QY 292 GTGAGGAGCAGATTTCAGAGGACAGTCCACCTCCGGGAACATGCAGGCAGTTGAACTG 351
Db 121 GTGAGGAGCAGATTTCAGAGGACAGTCCACCTCCGGGAACATGCAGGCAGTTGAACTG 180
QY 352 CTGCTGAGCACCTTGGAGAGGGAGTCTGGCACCTTGGTTGGACTCGGGAATTCGTGGAG 411
Db 181 CTGCTGAGCACCTTGGAGAGGGAGTCTGGCACCTTGGTTGGACTCGGGAATTCGTGGAG 240
QY 412 GCCCTCCGAGAACCCGGCAGCCCTCTGGCCGCCCGCTACATGAACCTGAGCTCAGGAC 471
Db 241 GCCCTCCGAGAACCCGGCAGCCCTCTGGCCGCCCGCTACATGAACCTGAGCTCAGGAC 300
QY 472 TTGCCCTCTCCATCGTTTGAAGACGCTCATGATGAATATCTCCAACCTGCTGAACCTCCTT 531
Db 301 TTGCCCTCTCCATCGTTTGAAGACGCTCATGATGAATATCTCCAACCTGCTGAACCTCCTT 360
QY 532 CAGCCCACTCTGGTGACAAAGCTTAGTTAGAGACGCTTTGGATAAGTGCATGGAGGAG 591
Db 361 CAGCCCACTCTGGTGACAAAGCTTAGTTAGAGACGCTTTGGATAAGTGCATGGAGGAG 420
QY 592 GAACTGTTGACAAATTGAAGACAGAAAAACCGGATTGCTGCTGCAGAAAAACAATGGAAATGAA 651
Db 421 GAACTGTTGACAAATTGAAGACAGAAAAACCGGATTGCTGCTGCAGAAAAACAATGGAAATGAA 480
QY 652 TCAGGTGTAAGAGAGCTACTAAAAAGGATTGTGCAGAAAGAAACTGGTTCTCTGCATTT 711
Db 481 TCAGGTGTAAGAGAGCTACTAAAAAGGATTGTGCAGAAAGAAACTGGTTCTCTGCATTT 540
QY 712 CTGAATGTTCTTCGTCAACACAGGAAACAATGAACCTTGTCGAAGAGTTAACAGGCTCTGAT 771
Db 541 CTGAATGTTCTTCGTCAACACAGGAAACAATGAACCTTGTCGAAGAGTTAACAGGCTCTGAT 600
QY 772 TGCTCAGAAAGCAATGCAGAGATTGAGAATTTATCACAAGTTGATGGTCC 821
Db 601 TGCTCAGAAAGCAATGCAGGTTATTGTAATTTTACTGAGGAAGATTCTTTC 650

RESULT 15
BX492926
LOCUS
DEFINITION
BX492926 DKFZp781C0628_r1 781 (synonym: hlcc4) Homo sapiens cDNA clone
ACCESSION
BX492926 DKFZp781C0628_5', mRNA sequence.
VERSION
BX492926.1 GI:32004887

KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 627)
AUTHORS Wambutt,R., Heubner,D., Mewes,H.W., Weil,B., Amid,C., Osanger,A.,
Fobo,G., Han,M. and Wiemann,S.
TITLE EST (Wambutt,R., Heubner,D., Mewes,H.W., Weil,B., Amid,C., et al.)
JOURNAL Unpublished (2003)
COMMENT Contact: MIPS

MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing
consortium of the German Genome Project.
No s1 sequence available.
This clone (DKFZp781C0628) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES
source
1..627
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp781C0628"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="781 (synonym: hlcc4)"
/note="Vector: pSport1_Sfi; Site_1: SfiIA; Site_2: SfiIB;
cDNA-collection"

ORIGIN

Query Match 18.6%; Score 625.4; DB 5; Length 627;
Best Local Similarity 99.8%; Pred. No. 2.1e-141;
Matches 626; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1372 GAAGTTGTCAAGTCCTGTGATATTATTATCAGTACAGCTCAAAATCCTTGAAAACTCCCTC 1431
|||||
Db 1 GAAGTTGTCAAGTCCTGTGATATTATTATCAGTACAGCTCAAAATCCTTGAAAACTCCCTC 60

QY 1432 TTAACCTTGGAAAAATGGAGAAGATGCTGGTGTTCATTTGTCAGACTTTTCCCTCATTATC 1491
|||||
Db 61 TTAACCTTGGAAAAATGGAGAAGATGCTGGTGTTCATTTGTCAGACTTTTCCCTCATTATC 120

QY 1492 ATTGATGAATGTCATCACACCAACAAGAAGCAGTGTAATAAATCAATCATGAGGCATTAT 1551
|||||
Db 121 ATTGATGAATGTCATCACACCAACAAGAAGCAGTGTAATAAATCAATCATGAGGCATTAT 180

QY 1552 TTGATGCAGAAGTTGAAAAACAATAGACTCAAGAAAGAAAAAACCAGTGATTCCTCCTT 1611
|||||
Db 181 TTGATGCAGAAGTTGAAAAACAATAGACTCAAGAAAGAAAAAACCAGTGATTCCTCCTT 240

QY 1612 CCTCAGATACTGGGACTAACAGCTTCACCTGGTGTGGAGGGGCCACGAAGCAAGCCAAA 1671
|||||
Db 241 CCTCAGATACTGGGACTAACAGCTTCACCTGGTGTGGAGGGGCCACGAAGCAAGCCAAA 300

QY 1672 GCTGAAGAACACATTTTAAAACTATGTGCCAATCTTGATGCATTTACTATTAAAACTGTT 1731
|||||
Db 301 GCTGAAGAACACATTTTAAAACTATGTGCCAATCTTGATGCATTTACTATTAAAACTGTT 360

QY 1732 AAAGAAAACCTTGATCAACTGAAAAACCAAAATACAGGAGCCATGCAAGAAGTTTGCCATT 1791
|||||
Db 361 AAAGAAAACCTTGATCAACTGAAAAACCAAAATACAGGAGCCATGCAAGAAGTTTGCCATT 420

QY 1792 GCAGATGCAACCAAGAGAAGATCCATTAAAGAGAAAACCTTAGAAAATAATGACAAGGATT 1851
|||||
Db 421 GCAGATGCAACCAAGAGAAGATCCATTAAAGAGAAAACCTTAGAAAATAATGACAAGGATT 480

QY 1852 CAAACTTATTGTCAAATGAGTCCCAATGTGAGATTTTGGAACTCAACCCCTATGAACAATGG 1911
|||||

Db 481 CAAACTTATTGTCAAATGAGTCCAATGTGAGATTTTGGAACTCAACCCCTATGAACAATGG 540
QY 1912 GCCATTCAAATGGAAAAAAGCTGCAAAAAAAGGAAATCGCAAAAGAACGTGTTTGTGCA 1971
|||||
Db 541 GCCATTCAAATGGAAAAAAGCTGCAAAAAAAGGAAATCGCAAAAGAACGTGTTTGTGCA 600
QY 1972 GAACATTGAGGAAGTACAATGAGGCC 1998
|||||
Db 601 GAACATTGAGGAAGTACAATGAGGCC 627

Search completed: January 27, 2005, 17:52:21
Job time : 9214 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 26, 2005, 19:38:26 ; Search time 165 Seconds
(without alignments)
2228.470 Million cell updates/sec

Title: US-09-515-363C-2
Perfect score: 5311
Sequence: 1 MSGYSTDENFRYLISCFRA.....LPITFPNLDYSECCLFSEDED 1025

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Sep04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				.%		Query		Match		Length		DB		ID		Description	
Result No.	Score	Match	Length	DB	ID	Match	Length	DB	ID	Match	Length	DB	ID	Match	Length	DB	ID
1	5311	100.0	1025	4	AAE10155	100.0	1025	4	AAE10155	100.0	1025	4	AAE10155	100.0	1025	4	AAE10155 Human mel
2	5311	100.0	1025	8	ADJ75424	100.0	1025	8	ADJ75424	100.0	1025	8	ADJ75424	100.0	1025	8	ADJ75424 Marker ge
3	5311	100.0	1025	8	ADN04880	100.0	1025	8	ADN04880	100.0	1025	8	ADN04880	100.0	1025	8	ADN04880 Antipsori
4	5285	99.5	1025	5	AAM47798	99.5	1025	5	AAM47798	99.5	1025	5	AAM47798	99.5	1025	5	AAM47798 Human RNA
5	5285	99.5	1025	7	ADC31794	99.5	1025	7	ADC31794	99.5	1025	7	ADC31794	99.5	1025	7	ADC31794 Human nov
6	4306	81.1	838	4	AAE10165	81.1	838	4	AAE10165	81.1	838	4	AAE10165	81.1	838	4	AAE10165 RNA helic
7	4194	79.0	1025	8	ADJ76213	79.0	1025	8	ADJ76213	79.0	1025	8	ADJ76213	79.0	1025	8	ADJ76213 Marker ge
8	2595.5	48.9	558	4	AAU23090	48.9	558	4	AAU23090	48.9	558	4	AAU23090	48.9	558	4	AAU23090 Novel hum
9	2456	46.2	468	6	ADA54899	46.2	468	6	ADA54899	46.2	468	6	ADA54899	46.2	468	6	ADA54899 Human pro
10	1860	35.0	417	4	AAU23647	35.0	417	4	AAU23647	35.0	417	4	AAU23647	35.0	417	4	AAU23647 Novel hum
11	1839	34.6	356	4	AAM40129	34.6	356	4	AAM40129	34.6	356	4	AAM40129	34.6	356	4	AAM40129 Human pol
12	1769	33.3	348	4	AAM41915	33.3	348	4	AAM41915	33.3	348	4	AAM41915	33.3	348	4	AAM41915 Human pol
13	1769	33.3	348	7	ADC33339	33.3	348	7	ADC33339	33.3	348	7	ADC33339	33.3	348	7	ADC33339 Human nov
14	1352	25.5	678	5	ABP69672	25.5	678	5	ABP69672	25.5	678	5	ABP69672	25.5	678	5	ABP69672 Human pol
15	1344	25.3	678	4	AAB93708	25.3	678	4	AAB93708	25.3	678	4	AAB93708	25.3	678	4	AAB93708 Human pro
16	1093	20.6	304	4	AAU00296	20.6	304	4	AAU00296	20.6	304	4	AAU00296	20.6	304	4	AAU00296 Interfero
17	1093	20.6	308	4	AAU00298	20.6	308	4	AAU00298	20.6	308	4	AAU00298	20.6	308	4	AAU00298 Interfero
18	1087.5	20.5	925	7	ADJ70225	20.5	925	7	ADJ70225	20.5	925	7	ADJ70225	20.5	925	7	ADJ70225 Human hea
19	1086.5	20.5	925	7	ADF76335	20.5	925	7	ADF76335	20.5	925	7	ADF76335	20.5	925	7	ADF76335 Novel hum
20	1086.5	20.5	925	8	ADJ75428	20.5	925	8	ADJ75428	20.5	925	8	ADJ75428	20.5	925	8	ADJ75428 Marker ge
21	1086.5	20.5	925	8	ADJ75495	20.5	925	8	ADJ75495	20.5	925	8	ADJ75495	20.5	925	8	ADJ75495 Marker ge
22	1086.5	20.5	925	8	ADN04860	20.5	925	8	ADN04860	20.5	925	8	ADN04860	20.5	925	8	ADN04860 Antipsori
23	1040.5	19.6	487	7	ADM05830	19.6	487	7	ADM05830	19.6	487	7	ADM05830	19.6	487	7	ADM05830 Human pro
24	861	16.2	267	4	AAU00297	16.2	267	4	AAU00297	16.2	267	4	AAU00297	16.2	267	4	AAU00297 Interfero
25	861	16.2	271	4	AAU00299	16.2	271	4	AAU00299	16.2	271	4	AAU00299	16.2	271	4	AAU00299 Interfero

26	808	15.2	448	4	AAU23654	Aau23654 Novel hum
27	783	14.7	166	4	ABB42219	Abb42219 Peptide #
28	783	14.7	166	4	AAM36026	Aam36026 Peptide #
29	783	14.7	166	4	AAM75915	Aam75915 Human bon
30	783	14.7	166	4	AAM63104	Aam63104 Human bra
31	782	14.7	154	4	AAG74314	Aag74314 Human col
32	782	14.7	154	5	ABP41112	Abp41112 Human ova
33	666.5	12.5	357	3	AAB42981	Aab42981 Human ORF
34	563.5	10.6	447	5	ABP41171	Abp41171 Human ova
35	558.5	10.5	447	4	AAB95485	Aab95485 Human pro
36	521	9.8	312	4	AAU23099	Aau23099 Novel hum
37	488	9.2	92	4	AAM14138	Aam14138 Peptide #
38	488	9.2	92	4	ABB33083	Abb33083 Peptide #
39	488	9.2	92	4	AAM26544	Aam26544 Peptide #
40	488	9.2	92	4	ABB27911	Abb27911 Human pep
41	488	9.2	92	4	ABBI8551	Abbi8551 Protein #
42	488	9.2	92	4	AAM66268	Aam66268 Human bon
43	488	9.2	92	4	AAM53880	Aam53880 Human bra
44	488	9.2	92	4	ABG47932	Abg47932 Human liv
45	488	9.2	92	4	AAM01876	Aam01876 Peptide #

ALIGNMENTS

RESULT 1
AAE10155
ID AAE10155 standard; protein; 1025 AA.
XX
AC AAE10155;
XX
DT 29-NOV-2001 (first entry)
XX
DE Human melanoma differentiation associated (Mda)-5 protein.
XX
KW Human; melanoma differentiation associated gene; Mda-5; interferon; IFN;
KW RNA helicase motif; caspase recruitment domain; CARD; therapy; melanoma;
KW neuroblastoma; astrocytoma; glioblastoma; multiforme; cancer; cervical;
KW breast; colon; prostate; osteosarcoma; chondrosarcoma; systemic toxicity;
KW central nervous system; cytostatic; apoptosis.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Region 125..174
FT /note= "Reveals sequence homology with other CARD
FT proteins"
FT Domain 331..336
FT /note= "ATPase A motif"
FT Domain 443..446
FT /note= "ATPase B motif"
FT Region 722..823
FT /note= "Reveals significant homology to RNA helicase C-
FT terminal conserved domain"
XX
PN WO200164707-A1.
XX
PD 07-SEP-2001.
XX
PF 28-FEB-2001; 2001WO-US0006960.
XX
PR 29-FEB-2000; 2000US-00515363.
XX
PA (UYCO) UNIV COLUMBIA NEW YORK.
XX
PI Fisher PB, Kang D, Gopalkrishnan RV;
XX
DR WPI; 2001-565494/63.
DR N-PSDB; AAD17203.
XX
PT Nucleic acid sequences encoding a Melanoma Differentiation Associated
PT Gene useful for cancer cell growth suppression, apoptosis and anti-viral
PT activity.

CC stimulated with interleukin-13. Also described: (1) a reagent (I) for
CC testing for bronchial asthma or chronic obstructive pulmonary disease;
CC (2) a kit for screening for a candidate compound for a therapeutic agent
CC to treat bronchial asthma or chronic obstructive pulmonary disease; (3)
CC an animal model for bronchial asthma or chronic obstructive pulmonary
CC disease; (4) an inducer that induces bronchial asthma in a mouse; (5) a
CC method for producing an animal model for bronchial asthma or chronic
CC obstructive pulmonary disease; (6) a therapeutic agent for bronchial
CC asthma or chronic obstructive pulmonary disease, comprising the compound,
CC a marker gene or an antisense nucleic acid corresponding to a portion of
CC the marker gene, a ribozyme, a polynucleotide that suppresses the
CC expression of the gene through an RNAi effect or an antibody recognising
CC a protein encoded by a marker gene; and (7) a DNA chip for testing for
CC bronchial asthma or a chronic obstructive pulmonary disease, on which a
CC probe has been immobilised to assay a marker gene. (I) has respiratory
CC and antiasthmatic activities, and can be used in gene therapy. The method
CC is useful for testing for or screening for a therapeutic agent for
CC bronchial asthma or chronic obstructive pulmonary disease. The present
CC sequence is used in the exemplification of the present invention.
XX
SQ Sequence 1025 AA;

Query Match 100.0%; Score 5311; DB 8; Length 1025;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1025; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSGYSTDENFRYLISCFRVRKMYIQVEPVLDTLFLPAEYKEIQIORTVATSGNQAVE 60
Db 1 MSGYSTDENFRYLISCFRVRKMYIQVEPVLDTLFLPAEYKEIQIORTVATSGNQAVE 60
QY 61 LLLSTLEKGVHGLGWTRFVEALRRITGSPLAARYMNPDLTLPSPSFENAHDEYLQLNL 120
Db 61 LLLSTLEKGVHGLGWTRFVEALRRITGSPLAARYMNPDLTLPSPSFENAHDEYLQLNL 120
QY 121 LQPTLVKLLVRDVLDKCMEEEELLTIEDNRNRIAAEENNGESGVRELLKRIVQKNWESA 180
Db 121 LQPTLVKLLVRDVLDKCMEEEELLTIEDNRNRIAAEENNGESGVRELLKRIVQKNWESA 180
QY 181 FLNVLRTGNNELVQELTGSDCESNAEINLSQVDPQVEBQLLSTTVQPNLEKEVWGM 240
Db 181 FLNVLRTGNNELVQELTGSDCESNAEINLSQVDPQVEBQLLSTTVQPNLEKEVWGM 240
QY 241 ENNSSESSFADSSVSVESDTSLAEGSVSCLDESGLHNSNMGSDSGTMGSDSDEENVAARA 300
Db 241 ENNSSESSFADSSVSVESDTSLAEGSVSCLDESGLHNSNMGSDSGTMGSDSDEENVAARA 300
QY 301 SPEPELQLRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKDHLDKKKASEPGKVI 360
Db 301 SPEPELQLRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKDHLDKKKASEPGKVI 360
QY 361 VLNVKLVLLVEQLPRKEFPFLKKWYRVIGLSGDTQLKISFPEVVKSCDIIISTAQILENS 420
Db 361 VLNVKLVLLVEQLPRKEFPFLKKWYRVIGLSGDTQLKISFPEVVKSCDIIISTAQILENS 420
QY 421 LLNLENGEDAGVQLSDFSLIIIDECHHTNKEAVYNNIMRHYLMQKLKNRLKKNKPVIP 480
Db 421 LLNLENGEDAGVQLSDFSLIIIDECHHTNKEAVYNNIMRHYLMQKLKNRLKKNKPVIP 480
QY 481 LPQILGLTASPGVGATKQAKAEEHILKLCANLDAFTIKTVKENLDQLKNQIQEPCCKFA 540
Db 481 LPQILGLTASPGVGATKQAKAEEHILKLCANLDAFTIKTVKENLDQLKNQIQEPCCKFA 540
QY 541 IADATREDPFKEKLEIMTRIOTYQCMSPMSDFGTQPYEQWAIQMEKKAAGKGNKERV 600
Db 541 IADATREDPFKEKLEIMTRIOTYQCMSPMSDFGTQPYEQWAIQMEKKAAGKGNKERV 600
QY 601 AEHLRKYNEALQINDTIRMIDAYTHLETIFYNEEKDKKFAVIEDSDSDEGDDDEYCDGDEDE 660
Db 601 AEHLRKYNEALQINDTIRMIDAYTHLETIFYNEEKDKKFAVIEDSDSDEGDDDEYCDGDEDE 660
QY 661 DDLKKPLKLDDETRFLMTLFFENNNKMLKRLAENPEYENKLTCLRNTIMEQYTRTESAR 720
Db 661 DDLKKPLKLDDETRFLMTLFFENNNKMLKRLAENPEYENKLTCLRNTIMEQYTRTESAR 720

QY 721 GIIFTKTRQSAYALSQWITENEKFAEVGVKAHHLIGAGHSSEFKPMTQNEQKEVISKERT 780
Db 721 GIIFTKTRQSAYALSQWITENEKFAEVGVKAHHLIGAGHSSEFKPMTQNEQKEVISKERT 780
QY 781 GKINLLIATTVAEGLDIKECNIVIRYGLVTNEIAMVQARGRARADESTYVLVAHSGSGV 840
Db 781 GKINLLIATTVAEGLDIKECNIVIRYGLVTNEIAMVQARGRARADESTYVLVAHSGSGV 840
QY 841 IEHETVNDPREKMMYKAHCVQNMKPEEYAHKILELQMSIMEKMKTKRNIKHYKNP 900
Db 841 IEHETVNDPREKMMYKAHCVQNMKPEEYAHKILELQMSIMEKMKTKRNIKHYKNP 900
QY 901 SLITFLCKNCVSLACSGEDIHVIEKMHVNMTPEFKELYIVRENKALQKKCADIQINGEI 960
Db 901 SLITFLCKNCVSLACSGEDIHVIEKMHVNMTPEFKELYIVRENKALQKKCADIQINGEI 960
QY 961 ICKCGQAWGTMVMVHKGLDLPCLKIRNFVVVKNNSTKKQYKKWVELPITFPNLDYSECCL 1020
Db 961 ICKCGQAWGTMVMVHKGLDLPCLKIRNFVVVKNNSTKKQYKKWVELPITFPNLDYSECCL 1020
QY 1021 FSDED 1025
Db 1021 FSDED 1025

RESULT 3
ADN04880
ID ADN04880 standard; protein; 1025 AA.

XX ADN04880;
XX
DT 01-JUL-2004 (first entry)
XX
DE Antipsoriatic protein sequence #620.
XX antipsoriatic; gene therapy; psoriasis; diagnosis.
XX Homo sapiens.
XX WO2004028479-A2.
XX
PD 08-APR-2004.
XX
PF 25-SEP-2003; 2003WO-US030907.
XX
PR 25-SEP-2002; 2002US-0414006P.
XX
PA (GETH) GENENTECH INC.
XX

PI Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;
PI Wu TD;
XX
DR WPI; 2004-305105/28.
DR N-PSDB; ADN04879.
XX
PT New PRO nucleic acid or polypeptide, useful for preparing a
PT pharmaceutical composition for diagnosing or treating psoriasis in a
XX mammal.
PS Claim 9; SEQ ID NO 1274; 3069pp; English.
XX
CC The invention relates to novel polynucleotide and polypeptides for
CC treating psoriasis or a sequence having at least 80% identity to the
CC above sequences. The nucleic acid is useful for preparing a composition
CC for diagnosing or treating psoriasis in a mammal. This sequence
CC corresponds to one of the polypeptides of the invention.
SQ Sequence 1025 AA;

Query Match 100.0%; Score 5311; DB 8; Length 1025;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1025; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MSGYSTDENFRYLISCFRARKVMIQVEPVLDYLTFLPAEVKEIQRTVATSGNMQAVE	60
Db	1	MSGYSTDENFRYLISCFRARKVMIQVEPVLDYLTFLPAEVKEIQRTVATSGNMQAVE	60
QY	61	LLSTLEKGVWHLGWTREFVEALRRTGSPLAARYMNPDLTDLSPSFENAHDEYLQLLNL	120
Db	61	LLSTLEKGVWHLGWTREFVEALRRTGSPLAARYMNPDLTDLSPSFENAHDEYLQLLNL	120
QY	121	LQPTLVKLLVRDVLKCMEEELLTIEDNRNIAAENNGNESGVRELLKRIVQKENWFS	180
Db	121	LQPTLVKLLVRDVLKCMEEELLTIEDNRNIAAENNGNESGVRELLKRIVQKENWFS	180
QY	181	FLNVLRTQGNNELVQELTGSDCSSESNAEIEENLSQVDGPQVEEQLLSTTVQPNLEKEVWGM	240
Db	181	FLNVLRTQGNNELVQELTGSDCSSESNAEIEENLSQVDGPQVEEQLLSTTVQPNLEKEVWGM	240
QY	241	ENNSSESFADSSVSVSESDTSLAEGSVCLDESGLHNSNMGSDSGTGMGSDSDSEENVAARA	300
Db	241	ENNSSESFADSSVSVSESDTSLAEGSVCLDESGLHNSNMGSDSGTGMGSDSDSEENVAARA	300
QY	301	SPEPELQLRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKDHLDKKKKASEPGKVI	360
Db	301	SPEPELQLRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKDHLDKKKKASEPGKVI	360
QY	361	VLNVKLLVEQLFRKEFPFLKKYRVIGLSGDTQLKISFPEVVKSCDIIISTAQILENS	420
Db	361	VLNVKLLVEQLFRKEFPFLKKYRVIGLSGDTQLKISFPEVVKSCDIIISTAQILENS	420
QY	421	LLNLENGEDAGVQLSDFSLIIIDECHHTNKEAVYNNIMRHYLMQKLKNNRLKKNKPVIP	480
Db	421	LLNLENGEDAGVQLSDFSLIIIDECHHTNKEAVYNNIMRHYLMQKLKNNRLKKNKPVIP	480
QY	481	LPQILGLTASPGVGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQIQEPCKKFA	540
Db	481	LPQILGLTASPGVGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQIQEPCKKFA	540
QY	541	IADATREDPFKEKLEIMTRIQTQYCMSPMSDFGTQPYEQWAIQMEKKAACKGNRKERV	600
Db	541	IADATREDPFKEKLEIMTRIQTQYCMSPMSDFGTQPYEQWAIQMEKKAACKGNRKERV	600
QY	601	AEHLRKYNEALQINDTIRMIDAYTHLETFFYNEEKDKKFAVIEDSDDEGDDDEYCDGDEDE	660
Db	601	AEHLRKYNEALQINDTIRMIDAYTHLETFFYNEEKDKKFAVIEDSDDEGDDDEYCDGDEDE	660
QY	661	DDLKKPLKLDLDRFLMTLFFENNMKMLKRLAENPEYENEKLTKLRLNTIMEQYTRTESAR	720
Db	661	DDLKKPLKLDLDRFLMTLFFENNMKMLKRLAENPEYENEKLTKLRLNTIMEQYTRTESAR	720
QY	721	GIIFTKTRQAYALSQWITENEKFAEVGVKAHHLIGAGHSSEFFKPMTQNEQKEVISKERT	780
Db	721	GIIFTKTRQAYALSQWITENEKFAEVGVKAHHLIGAGHSSEFFKPMTQNEQKEVISKERT	780
QY	781	GKINLLIATTVAEEGLDIKECNIVIRYGLVTNEIAMVQARGRARADESTYVLVAHSGSGV	840
Db	781	GKINLLIATTVAEEGLDIKECNIVIRYGLVTNEIAMVQARGRARADESTYVLVAHSGSGV	840
QY	841	IEHETVNDPREKMYKAIHCQVQNMKPEEYAHKILELQMQSIMEKKMKTARNIAKHYNKP	900
Db	841	IEHETVNDPREKMYKAIHCQVQNMKPEEYAHKILELQMQSIMEKKMKTARNIAKHYNKP	900
QY	901	SLITFLCKNCVSLACSGEDIHVEIKMHVNMTPPEFKELYIVRENKALQKCADYQINGEI	960
Db	901	SLITFLCKNCVSLACSGEDIHVEIKMHVNMTPPEFKELYIVRENKALQKCADYQINGEI	960
QY	961	ICKCGQAWGTMVHKGLDLPCLKIRNFVVFKNNSKKQYKKWVELPITFPNLDYSECCL	1020
Db	961	ICKCGQAWGTMVHKGLDLPCLKIRNFVVFKNNSKKQYKKWVELPITFPNLDYSECCL	1020
QY	1021	FSDED 1025	
Db	1021	FSDED 1025	

RESULT 4			
AAM47798			
ID	AAM47798	standard; protein; 1025 AA.	
XX			
AC	AAM47798;		
DT	05-MAR-2002	(first entry)	
XX			
DE	Human RNA helicase RH116.		
XX			
KW	Human; RH116; RNA helicase; cytostatic; virucide; anti-HIV;		
KW	immunosuppressive; immunostimulatory; antirheumatic; antiarthritic;		
KW	antiartherosclerotic; osteopathic; antidiabetic; hepatotropic;		
KW	antiinflammatory; cancer; infection; HIV; hepatitis; genetic disease;		
KW	autoimmune disease; graft rejection; vaccine.		
XX			
OS	Homo sapiens.		
XX			
PN	WO200185955-A1.		
XX			
PD	15-NOV-2001.		
XX			
PF	11-MAY-2001; 2001WO-FR001441.		
XX			
PR	11-MAY-2000; 2000FR-00006030.		
XX			
PA	(ISTA-) ISTAC.		
PA	(INSP) INST PASTEUR LILLE.		
XX			
XX	Bahr G, Cocude C, Capron A;		
PI	WPI; 2002-082898/11.		
DR	N-PSDB; ABA04908.		
DR			
XX			
PT	New polypeptide, useful for treating and diagnosing cancer or		
PT	inflammation, and drug screening, comprises a human polynucleotide		
PT	homologous to RNA helicase.		
XX			
PS	Claim 1; Page 89-93; 114pp; French.		
XX			
CC	The present sequence is the protein sequence for human RH116. RH116 is a		
CC	116kDa protein and has homology to RNA helicases (DEXH box). RH116 and		
CC	its coding sequence are useful for treating cancer; acute or chronic		
CC	infections (especially by HIV or hepatitis B or C); inherited genetic		
CC	diseases; (auto)immune diseases (particularly rheumatism, arthritis,		
CC	arteriosclerosis, osteoporosis and diabetes, but many others listed) and		
CC	to prevent graft rejection. RH116 and its coding sequence are also useful		
CC	for inducing, or increasing, the immune response to a vaccine		
XX			
SQ	Sequence 1025 AA;		
Query Match 99.5%; Score 5285; DB 5; Length 1025;			
Best Local Similarity 99.5%; Pred. No. 0;			
Matches 1020; Conservative 2; Mismatches 3; Indels 0; Gaps 0;			
QY	1	MSGYSTDENFRYLISCFRARKVMIQVEPVLDYLTFLPAEVKEIQRTVATSGNMQAVE	60
Db	1	MSGYSTDENFRYLISCFRARKVMIQVEPVLDYLTFLPAEVKEIQRTVATSGNMQAVE	60
QY	61	LLSTLEKGVWHLGWTREFVEALRRTGSPLAARYMNPDLTDLSPSFENAHDEYLQLLNL	120
Db	61	LLSTLEKGVWHLGWTREFVEALRRTGSPLAARYMNPDLTDLSPSFENAHDEYLQLLNL	120
QY	121	LQPTLVKLLVRDVLKCMEEELLTIEDNRNIAAENNGNESGVRELLKRIVQKENWFS	180
Db	121	LQPTLVKLLVRDVLKCMEEELLTIEDNRNIAAENNGNESGVRELLKRIVQKENWFS	180
QY	181	FLNVLRTQGNNELVQELTGSDCSSESNAEIEENLSQVDGPQVEEQLLSTTVQPNLEKEVWGM	240
Db	181	FLNVLRTQGNNELVQELTGSDCSSESNAEIEENLSQVDGPQVEEQLLSTTVQPNLEKEVWGM	240
QY	241	ENNSSESFADSSVSVSESDTSLAEGSVCLDESGLHNSNMGSDSGTGMGSDSDSEENVAARA	300

Db 241 ENNSSESFADSSVVSSESDTSLAEGSVSCLDESGLHNSNMGSDSGTMGSDSDEENVAARA 300
QY 301 SPEPELQLRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKDHLDKXKASEPGKVI 360
Db 301 SPEPELQLRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKDHLDKXKASEPGKVI 360
QY 361 VLVNKVLLVEQLFRKEFQPFLLKKWYRVIGLSGDTQLKISFPPEVVKSCDIIISTAQILENS 420
Db 361 VLVNKVLLVEQLFRKEFQPFLLKKWYRVIGLSGDTQLKISFPPEVVKSCDIIISTAQILENS 420
QY 421 LNLNLENGEDAGVQLSDFSLIIIDECHHTNKEAVYNNIMRHYLMQKLNRLKKNKPVIP 480
Db 421 LNLNLENGEDAGVQLSDFSLIIIDECHHTNKEAVYNNIMRHYLMQKLNRLKKNKPVIP 480
QY 481 LPQILGLTASPGVGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQIQEPCKKPA 540
Db 481 LPQILGLTASPGVGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQIQEPCKKPA 540
QY 541 IADATREDPFKEKLEIMTRIQTQYQMSPMDSDFGTQPYEQWATQMEKKAAGNKRKERV 600
Db 541 IADATREDPFKEKLEIMTRIQTQYQMSPMDSDFGTQPYEQWATQMEKKAAGNKRKERV 600
QY 601 AEHLRKYNEALQINDTIRMIDAYTHLETFYNEEKDKKFAVIEDDSDEGGDDDEYCDGDEDE 660
Db 601 AEHLRKYNEALQINDTIRMIDAYTHLETFYNEEKDKKFAVIEDDSDEGGDDDEYCDGDEDE 660
QY 661 DDLKKPLKLDDETRFLMTLFFENNKMMLKRLAENPEYENEKLTCLRNTIMEQYTRTEESAR 720
Db 661 DDLKKPLKLDDETRFLMTLFFENNKMMLKRLAENPEYENEKLTCLRNTIMEQYTRTEESAR 720
QY 721 GIIFTKTRQSAVALSQWITENEKFAEVGVKAHHLIGAGHSSEFPKPMTQNEQKEVSKFRT 780
Db 721 GIIFTKTRQSAVALSQWITENEKFAEVGVKAHHLIGAGHSSEFPKPMTQNEQKEVSKFRT 780
QY 781 GKINLLIATTVAEEGLDIKECNIVIRYGLVTNEIAMVQARGARADESTYVLVAHSGSGV 840
Db 781 GKINLLIATTVAEEGLDIKECNIVIRYGLVTNEIAMVQARGARADESTYVLVAHSGSGV 840
QY 841 IEHETVNDFREKMYKAIHCVQNMKPEYAHKILELQMQSIMEKMKTKRNIKHYKNP 900
Db 841 IERETVNDFREKMYKAIHCVQNMKPEYAHKILELQMQSIMEKMKTKRNIKHYKNP 900
QY 901 SLITFLCKNCSVLACSGEDIHVIEKMHVNMTPEFKELYIVRENKALQKCCADYQINGEI 960
Db 901 SLITFLCKNCSVLACSGEDIHVIEKMHVNMTPEFKELYIVRENKALQKCCADYQINGEI 960
QY 961 ICKGQAWGTMVHKGLDLPCLKIRNFVVFKNNSTKKQYKQWVLPITFPNLDYSECCL 1020
Db 961 ICKGQAWGTMVHKGLDLPCLKIRNFVVFKNNSTKKQYKQWVLPITFPNLDYSECCL 1020
QY 1021 FSDED 1025
Db 1021 FSDED 1025

RESULT 5

ADC31794

ID ADC31794 standard; protein; 1025 AA.

XX ADC31794;

AC ADC31794;

XX 18-DEC-2003 (first entry)

DT Human novel polypeptide sequence, SEQ ID NO:1876.

XX Human; diagnostic; drug screening; forensics; gene mapping;

DE Human novel polypeptide sequence, SEQ ID NO:1876.

XX Human; diagnostic; drug screening; forensics; gene mapping;

KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;

KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;

KW ulcers; osteoporosis; autoimmune disease; cancer;

KW molecular weight marker; food supplement; antiparkinsonian; nootropic;

KW neuroprotective; antianaemic; anticoagulant; thrombolytic; vulnerary;

KW antiulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;

gene therapy.
XX Homo sapiens.
OS WO2003029271-A2.
PN 10-APR-2003.
XX 24-SEP-2002; 2002WO-US030474.
XX 24-SEP-2001; 2001US-0324631P.
PR (HYSE-) HYSEQ INC.
XX Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;
PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
PI Haley-Vicente D, Drmanac RT;
XX WPI; 2003-371981/35.
DR N-PSDB; ADC30823.
XX
PT New polynucleotide and polypeptide useful for diagnosing, preventing or
PT treating conditions such as neurodegenerative diseases, anemias, platelet
PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
PT cancer.
XX Claim 20; SEQ ID NO 1876; 1185pp; English.
PS
XX The invention relates to 971 novel human cDNA sequences (ADC29919-
CC ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The
CC invention also relates to nucleic acid sequences over 99% identical with
CC the novel human cDNAs. The invention additionally encompasses expression
CC vectors and host cells comprising a nucleic acid of the invention; the
CC recombinant production of a polypeptide of the invention; an antibody
CC against a polypeptide of the invention; a method of detecting
CC polynucleotides or polypeptides of the invention; and methods of
CC identifying a compound which binds to a polypeptide of the invention. The
CC invention further discloses methods of preventing, treating or
CC ameliorating a medical condition; kits comprising polynucleotide probes
CC and/or monoclonal antibodies for carrying out the methods of the
CC invention; methods for the identification of compounds that modulate the
CC expression or activity of the polynucleotide and/or polypeptide; and 767
CC contig sequences corresponding to the cDNA sequences of the invention
CC (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628
CC -ADC33394). The nucleic acids and polypeptides of the invention are
CC useful in diagnostics, drug screening, forensics, gene mapping, in the
CC identification of mutations responsible for genetic disorders or other
CC traits, for assessing biodiversity, and in producing many other types of
CC data and products dependent on DNA and amino acid sequences. They are
CC also used for treating diseases such as Parkinson's disease, Alzheimer's
CC disease and other neurodegenerative diseases, anaemia, platelet
CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
CC cancer. The nucleic acids may also be used as hybridisation probes or
CC primers, and in the recombinant production of a protein. The polypeptides
CC are also useful in generating antibodies, as molecular weight markers,
CC and as food supplements. The present sequence represents a specifically
CC claimed human polypeptide sequence of the invention. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1025 AA;

Query Match 99.5%; Score 5285; DB 7; Length 1025;

Best Local Similarity 99.5%; Pred. No. 0;

Mismatches 1020; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MSGYSTDENFRYLISCFRARKVMYIQVEPVLDTLFLPAEVKEIQRTVATSGNMQAVE 60

Db 1 MSGYSTDENFRYLISCFRARKVMYIQVEPVLDTLFLPAEVKEIQRTVATSGNMQAVE 60

QY 61 LLLSTLEKGVHGLGWTREFVEALRRTGSPLAARYMNPDLTLPSPSFENAHDEYLQLNL 120

|||||

Db 61 LLLSTLEKGVWHLGWTRFVEALRRRTGSPLAARYMNPDLTDLSPSPENAHDEYLQLNL 120

QY 121 LQPTLVLDKLLVRDVLKCMEEELLTIEDNRNIAAENNGNESGVRELLKRIVQENWFSA 180

Db 121 LQPTLVLDKLLVRDVLKCMEEELLTIEDNRNIAAENNGNESGVRELLKRIVQENWFSA 180

QY 181 FLNVLROTGNNELVQELTGSDCSESNAEINLSQVDPQVVEQLLSTTVQPNLEKEVWGM 240

Db 181 FLNVLROTGNNELVQELTGSDCSESNAEINLSQVDPQVVEQLLSTTVQPNLEKEVWGM 240

QY 241 ENNSSESSPADSSVSVSESDTSLAEGSVSCLDESIGHNSNMGSDSGTSGSDSDSEENVAARA 300

Db 241 ENNSSESSPADSSVSVSESDTSLAEGSVSCLDESIGHNSNMGSDSGTSGSDSDSEENVAARA 300

QY 301 SPEPELQRPYQMEVAQPALEGNIIICLPTGSGKTRVAVYIAKDLKDKKASEPGKVI 360

Db 301 SPEPELQRPYQMEVAQPALEGNIIICLPTGSGKTRVAVYIAKDLKDKKASEPGKVI 360

QY 361 VLVNKVLLVEQLFRKEFPFLKKWYRVIGLSGDTQLKISFPVVKSCDIIISTAQILENS 420

Db 361 VLVNKVLLVEQLFRKEFPFLKKWYRVIGLSGDTQLKISFPVVKSCDIIISTAQILENS 420

QY 421 LLNLENGEDAGVQLSDFSLIIIDECHHTNKAEVYNNIMRHLYLMQKLKNNRLKKNKPVIP 480

Db 421 LLNLENGEDAGVQLSDFSLIIIDECHHTNKAEVYNNIMRHLYLMQKLKNNRLKKNKPVIP 480

QY 481 LPQILGLTASPGVGATKQAKAEHILKLCANLDAFTTKTVKENLDQLKNOIQEPCKKFA 540

Db 481 LPQILGLTASPGVGATKQAKAEHILKLCANLDAFTTKTVKENLDQLKNOIQEPCKKFA 540

QY 541 IADATREDPFKEKLEIMTRIQTVCQSPMSDFGTQPYEQWAIQMEKKAAGKGNRKERV 600

Db 541 IADATREDPFKEKLEIMTRIQTVCQSPMSDFGTQPYEQWAIQMEKKAAGKGNRKERV 600

QY 601 AEHLRKYNEALQINDTIRMIDAYTHLETFYNEEKDKKFAVIEDDSDEGGDDCYDGEDE 660

Db 601 AEHLRKYNEALQINDTIRMIDAYTHLETFYNEEKDKKFAVIEDDSDEGGDDCYDGEDE 660

QY 661 DDLKKPLKLDLDRFLMTLFFENNMKRLAENPEYENNEKLTCLRNTIMEQYTRTEESAR 720

Db 661 DDLKKPLKLDLDRFLMTLFFENNMKRLAENPEYENNEKLTCLRNTIMEQYTRTEESAR 720

QY 721 GIIFTKTQSAVALSQWITENEKFAEVGVKAHHLIGAGHSSEFKPMTQNEQKEVISKFT 780

Db 721 GIIFTKTQSAVALSQWITENEKFAEVGVKAHHLIGAGHSSEFKPMTQNEQKEVISKFT 780

QY 781 GKINLLIATTVAEEGLDIKECNIVIRYGLVTNEIAMVQARGARADESTYVLVAHSGSGV 840

Db 781 GKINLLIATTVAEEGLDIKECNIVIRYGLVTNEIAMVQARGARADESTYVLVAHSGSGV 840

QY 841 IEHETVNDFREKMYKAIHCVQNMKPEEYAHKILELQMQSIMEKKMKTGRNIAKHYKNP 900

Db 841 IERETVNDFREKMYKAIHCVQNMKPEEYAHKILELQMQSIMEKKMKTGRNIAKHYKNP 900

QY 901 SLITFLCKNCSVLACSGEDIHVIEKMHVNMTPEFKELYIVRENKALQKCCADYQINGEI 960

Db 901 SLITFLCKNCSVLACSGEDIHVIEKMHVNMTPEFKELYIVRENKALQKCCADYQINGEI 960

QY 961 ICKGQAGWTMMVHKGLDLPCLKIRNFVVVFKNNSTKKQYKWWVELPITFPNLDYSECL 1020

Db 961 ICKGQAGWTMMVHKGLDLPCLKIRNFVVVFKNNSTKKQYKWWVELPITFPNLDYSECL 1020

QY 1021 FSDED 1025

Db 1021 FSDED 1025

RESULT 6
AAE10165
ID AAE10165 standard; protein; 838 AA.
XX AC
XX AAE10165;

DT 29-NOV-2001 (first entry)

XX RNA helicase conserved motif of human Mda-5 protein.

DE Human; melanoma differentiation associated gene; Mda-5; interferon; IFN;

XX RNA helicase motif; caspase recruitment domain; CARD; therapy; melanoma;

KW neuroblastoma; astrocytoma; glioblastoma; multiforme; cancer; cervical;

KW breast; colon; prostate; osteosarcoma; chondrosarcoma; systemic toxicity;

KW central nervous system; cytostatic; apoptosis.

XX Homo sapiens.

OS WO200164707-A1.

XX 07-SEP-2001.

XX 28-FEB-2001; 2001WO-US006960.

XX 29-FEB-2000; 2000US-00515363.

XX (UYCO) UNIV COLUMBIA NEW YORK.

PA Fisher PB, Kang D, Gopalkrishnan RV;

XX WPI; 2001-565494/63.

XX Nucleic acid sequences encoding a Melanoma Differentiation Associated

PT Gene useful for cancer cell growth suppression, apoptosis and anti-viral

PT activity.

XX Example 1; Fig 1D; 152pp; English.

PS The present invention relates to an isolated nucleic acid encoding a

XX melanoma differentiation associated gene-5 (Mda-5) polypeptide. Mda-5

CC contains a caspase recruitment domain (CARD) and a RNA helicase motif.

CC Mda-5 is a novel interferon (IFN) inducible gene with structural

CC similarities to RNA helicases and CARD motif containing proteins. Mda-5

CC is induced during terminal differentiation in human melanoma cells

CC treated with the combination of recombinant fibroblast IFN and the

CC antileukemic compound mezerein (MEZ). Mda-5 is useful for identifying

CC compounds that may induce its expression. Mda-5 is useful for treating

CC cancer e.g., melanoma, neuroblastoma, astrocytoma, glioblastoma

CC multiforme, cervical cancer, breast cancer, colon cancer, prostate

CC cancer, osteosarcoma, chondrosarcoma and especially where the cancer is a

CC cancer of the central nervous system and apoptosis. The Mda-5 promoter

CC exhibits melanocyte tissue specificity and minimises systemic toxicity.

CC The present sequence is RNA helicase conserved motif of human Mda-5

CC protein

XX Sequence 838 AA;

SQ

Query Match 81.1%; Score 4306; DB 4; Length 838;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 838; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 LLLSTLEKGVWHLGWTRFVEALRRRTGSPLAARYMNPDLTDLSPSPENAHDEYLQLNL 120

Db 1 LLLSTLEKGVWHLGWTRFVEALRRRTGSPLAARYMNPDLTDLSPSPENAHDEYLQLNL 60

QY 121 LQPTLVLDKLLVRDVLKCMEEELLTIEDNRNIAAENNGNESGVRELLKRIVQENWFSA 180

Db 61 LQPTLVLDKLLVRDVLKCMEEELLTIEDNRNIAAENNGNESGVRELLKRIVQENWFSA 120

QY 181 FLNVLROTGNNELVQELTGSDCSESNAEINLSQVDPQVVEQLLSTTVQPNLEKEVWGM 240

Db 121 FLNVLROTGNNELVQELTGSDCSESNAEINLSQVDPQVVEQLLSTTVQPNLEKEVWGM 180

QY 241 ENNSSESSPADSSVSVSESDTSLAEGSVSCLDESIGHNSNMGSDSGTSGSDSDSEENVAARA 300

Db 181 ENNSSESSPADSSVSVSESDTSLAEGSVSCLDESIGHNSNMGSDSGTSGSDSDSEENVAARA 240

QY 301 SPEPELQRPYQMEVAQPALEGNIIICLPTGSGKTRVAVYIAKDLKDKKASEPGKVI 360

Db 241 SPEPELQLRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKHDLKDKKASEPGKVI 300

QY 361 VLVNKKVLLVEQLFRKEFPFLKKWYRVIGLSGDTQLKISFPEVVKSCDIIISTAQILENS 420

Db 301 VLVNKKVLLVEQLFRKEFPFLKKWYRVIGLSGDTQLKISFPEVVKSCDIIISTAQILENS 360

QY 421 LLNLENGEDAGVQLSDFSLIIIDECHHTNKEAVYNNIMRHYLMQKLNNRLKKNKPVIP 480

Db 361 LLNLENGEDAGVQLSDFSLIIIDECHHTNKEAVYNNIMRHYLMQKLNNRLKKNKPVIP 420

QY 481 LPQILGLTASPGVGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQIQEPCCKFA 540

Db 421 LPQILGLTASPGVGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQIQEPCCKFA 480

QY 541 IADATREDPFKEKLEIMTRIQTQYCMSPMSDFGTQPYEQWAIQMEKKAAGKGNRKERV 600

Db 481 IADATREDPFKEKLEIMTRIQTQYCMSPMSDFGTQPYEQWAIQMEKKAAGKGNRKERV 540

QY 601 AEHLRKYNEALQINDTIRMIDAYTHLETFYNEEKDKKFAVIEDSDSGDDEYCDGDEDE 660

Db 541 AEHLRKYNEALQINDTIRMIDAYTHLETFYNEEKDKKFAVIEDSDSGDDEYCDGDEDE 600

QY 661 DDLKKPLKLDLDRFLMTLFFENNMKLKRLAENPEYENEKLTCLRNTIMEQYTRTESAR 720

Db 601 DDLKKPLKLDLDRFLMTLFFENNMKLKRLAENPEYENEKLTCLRNTIMEQYTRTESAR 660

QY 721 GIIFTKTRQSAVALSQWITENEKFAEYGVKAHHLIGAGHSSEFKPMTQNEQKEVSKPRT 780

Db 661 GIIFTKTRQSAVALSQWITENEKFAEYGVKAHHLIGAGHSSEFKPMTQNEQKEVSKPRT 720

QY 781 GKINLLIATTVAEEGLDIKECNIVIRYGLVTNEIAMVQARGARADESTYVLVAHSGGV 840

Db 721 GKINLLIATTVAEEGLDIKECNIVIRYGLVTNEIAMVQARGARADESTYVLVAHSGGV 780

QY 841 IEHETVNDFREKMYKAIHCVQNMKPEYAHKILELQMQSIMEKKMKTNRNIAKHYN 898

Db 781 IEHETVNDFREKMYKAIHCVQNMKPEYAHKILELQMQSIMEKKMKTNRNIAKHYN 838

RESULT 7

ADJ76213

ID ADJ76213 standard; protein; 1025 AA.

XX

AC ADJ76213;

XX

DT 20-MAY-2004 (first entry)

XX

DE Marker gene related amino acid sequence SEQ ID NO:1465.

XX

KW bronchial asthma; chronic obstructive pulmonary disease;

KW respiratory epithelial cell; interleukin-13; respiratory; antiasthmatic;

KW gene therapy; marker.

XX

OS Mus musculus.

XX

PN EP1394274-A2.

XX

PD 03-MAR-2004.

XX

PF 04-AUG-2003; 2003EP-00254857.

XX

XX 06-AUG-2002; 2002JP-00229312.

PR 20-MAR-2003; 2003JP-00077212.

XX

PA (GENO-) GENOX RES INC.

XX

PI Ohtani N, Sugita Y, Yamaya M, Kubo H, Nagai H, Izuwara K;

XX

DR WPI; 2004-193155/19.

XX

PT Testing for bronchial asthma or chronic obstructive pulmonary disease by

PT comparing the expression level of a marker gene in a biological sample

PT from a subject with the expression level of the gene in a sample from a

PT healthy subject.

XX

PS Claim 16; SEQ ID NO 1465; 241pp; English.

XX

CC The present invention describes a method of testing for bronchial asthma or chronic obstructive pulmonary disease. The method comprises determining the expression level of a marker gene in a biological sample from a subject, comparing the expression level determined with the expression level of the marker gene in a biological sample from a healthy subject, and judging whether the subject has bronchial asthma or chronic obstructive pulmonary disease. The marker gene comprises: (a) a group of genes (S1) whose expression levels increase when respiratory epithelial cells are stimulated with interleukin-13; or (b) a group of genes (S2) whose expression levels decrease when respiratory epithelial cells are stimulated with interleukin-13. Also described: (1) a reagent (I) for testing for bronchial asthma or chronic obstructive pulmonary disease; (2) a kit for screening for a candidate compound for a therapeutic agent to treat bronchial asthma or chronic obstructive pulmonary disease; (3) an animal model for bronchial asthma or chronic obstructive pulmonary disease; (4) an inducer that induces bronchial asthma in a mouse; (5) a method for producing an animal model for bronchial asthma or chronic obstructive pulmonary disease; (6) a therapeutic agent for bronchial asthma or chronic obstructive pulmonary disease, comprising the compound, a marker gene or an antisense nucleic acid corresponding to a portion of the marker gene, a ribozyme, a polynucleotide that suppresses the expression of the gene through an RNAi effect or an antibody recognising a protein encoded by a marker gene; and (7) a DNA chip for testing for bronchial asthma or a chronic obstructive pulmonary disease, on which a probe has been immobilised to assay a marker gene. (I) has respiratory and antiasthmatic activities, and can be used in gene therapy. The method is useful for testing for or screening for a therapeutic agent for bronchial asthma or chronic obstructive pulmonary disease. The present sequence is used in the exemplification of the present invention.

XX

SQ Sequence 1025 AA;

Query Match 79.0%; Score 4194; DB 8; Length 1025;

Best Local Similarity 79.6%; Pred. No. 9.1e-316;

Matches 817; Conservative 75; Mismatches 132; Indels 2; Gaps 2;

QY 1 MSGYSTDENFRYLISCFRVRVMYIQVEPVLDTLFLPAEVKEIQRTVATSGNMQAVE 60

Db 1 MSIVCSAEDSFRNLILFFRPRLMYIQVEPVLDTLFLPAEVKEIQRTVATSGNMQAVE 60

QY 61 LLLSTLEKGVHGLWTRFVEALRRTGSPLAARYNPELTLPSFENAHDEYLQNL 120

Db 61 LLLSTLEKGVHGLWTRFVEALRRTGSPLAARYNPELTLPSFENAHDEYLQNL 120

QY 121 LQPTLVKLLVDRDLKCMEEELLTIEDNRNIAAENNGESGVRELLKRIVQKNWFS 180

Db 121 LQPTLVKLLVDRDLKCMEEELLTIEDNRNIAAENNGESGVRELLKRIVQKNWFS 180

QY 181 FLNVLRTGNNELVQELTGSDESNAEIEENLSQVDPQVEEQLLSTTVQPNLEKEVGM 240

Db 181 FLNVLRTGNNELVQELTGSDESNAEIEENLSQVDPQVEEQLLSTTVQPNLEKEVGM 240

QY 241 ENNSSESFPADSSVSESDTSLAEGSVCLDESGLHNSNMGSDSGTMGSDSDEENV-AAR 299

Db 241 DDILPEASCTDSSVTTESDTSIAEGSVSCFDESGLHNSNMGSDSGTMGSDSDEENV-AAR 299

QY 300 ASPEPELQLRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKHDLKDKKASEPGKV 359

Db 301 VSPEPELQLRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKHDLKDKKASEPGKV 360

QY 360 IVLVNKKVLLVEQLFRKEFPFLKKWYRVIGLSGDTQLKISFPEVVKSCDIIISTAQILEN 419

Db 361 IVLVNKKVLLVEQLFRKEFPFLKKWYRVIGLSGDTQLKISFPEVVKSCDIIISTAQILEN 420

QY 420 SLLNLENGEDAGVQLSDFSLIIIDECHHTNKEAVYNNIMRHYLMQKLNNRLKKNKPV 479

Db 421 SLLNLENGEDAGVQLSDFSLIIIDECHHTNKEAVYNNIMRHYLMQKLNNRLKKNKPV 480

QY 480 PLPQILGLTASPGVGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQIQEPCCKF 539

Db 481 PLPQILGLTASPGVGAAKQSEAEKHILNICANLDAFTTKVTVENLQGLKHQIKPECKKF 540
QY 540 AIADATREDPFKEKLLLEIMTRIQTYCQSPMSDFGTQPYEQWAIQMEKKAACKGNRKERV 599
Db 541 VIADDTRENPFKEKLLLEIMASIQTYCQSPMSDFGTQHYEQWAIQMEKKAADGNRKDRV 600
QY 600 CAEHLRKYNEALQINDTIRMIDAYTHLETFFYNEEKDKKFAVIEDDSDEGGDDEYCDGED 659
Db 601 CAEHLRKYNEALQINDTIRMIDAYSHLEAFYTDEKEKKFAVL-NDSDKDDDEASSCNDQL 659
QY 660 EDDLKPLKLDETDRFLMTLFFENNKMLKRLAENPEYENEKLTCLRNTIMEQYTRTEESA 719
Db 660 KGDVKSLKLDETDEFLMNLFFDNKKMLKCLAENPKYENEKLIKLRNTILEQFTRSESS 719
QY 720 RGIIFTKTROSAYALSQWITENEKFAEVGVKAHHLIGAGHSSEFKPMTQNEQKEVISKFR 779
Db 720 RGIIFTKTROSTYALSQWIMENAKFAEVGVKAHHLIGAGHSSEVKPMTQTEQKEVISKFR 779
QY 780 TGKINLLIATVAEEGLDIKECNIVIRYGLVTNEIAMVQARGRARADESTYVLVAHSGSG 839
Db 780 TGEINLLIATVAEEGLDIKECNIVIRYGLVTNEIAMVQARGRARADESTYVLVTSSGSG 839
QY 840 VIEHETVNDPREKQMYKAIHCVQNMKPEEYAHKILELQMSIMEKKMKTKRNI AKHYKNN 899
Db 840 VTEREIVNDPREKQMYKAINRVQNMKPEEYAHKILELQVQSILEKKMKVKRSIAKQYNDN 899
QY 900 PSLITFLCKNCVSLACSGEDIHVEIKMHVNMTPFEKELYIVRENKALQKKCADYQINGE 959
Db 900 PSLITLLCKNCMLVCSGENIHVEIKMHVNMTPFEKGLYIVRENKALQKKFADYQTNGE 959
QY 960 IICKCGQAWGTMVHKGLDLPCLKIRNFVVVFKNNSTKKQYKKWVELPITFPNLDYSECC 1019
Db 960 IICKCGQAWGTMVHKGLDLPCLKIRNFVVVFKNNSPKQYKKWVELPIRFPDLDYSEYC 1019
QY 1020 LFSDED 1025
Db 1020 LYSDED 1025

RESULT 8

AAU23090
ID AAU23090 standard; protein; 558 AA.
XX AC AAU23090;
XX DT 18-DEC-2001 (first entry)
XX DE Novel human enzyme polypeptide #176.
XX KW Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;
KW ligase; hyperproliferative disorder; immunodeficiency disorder;
KW autoimmune disorder; neurological disorder; metabolic disorder;
KW inflammatory disorder; cardiovascular disorder; reproductive disorder;
KW blood-related disorder; infectious disorder; cytostatic; anti arthritic;
KW nephrotropic; anticoagulant.
XX OS Homo sapiens.
XX PN WO200155301-A2.
XX PD 02-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US001239.
XX PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.

PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-465566/50.
DR N-PSDB; AAS40960.
XX

PT Novel polypeptides and polynucleotides useful for diagnosing, preventing,
PT treating neural, immune system, muscular, reproductive, pulmonary,
PT cardiovascular, renal, proliferative disorders and cancerous diseases.

XX
PS Claim 11; SEQ ID NO 1086; 1180pp; English.

XX
CC The present invention relates to the isolation of novel human enzyme
CC polypeptides, and the cDNA (AAS40785-AAS41684) and genomic sequences
CC encoding them. The enzyme polypeptides of the invention may comprise the
CC functional classes of oxidoreductases, transferases, hydrolases, lyases,

CC isomerases or ligases. The sequences of the invention are useful in the
CC diagnosis, treatment, prevention and/or prognosis of a wide range of
CC disorders, including hyperproliferative disorders (e.g. cancer),
CC immunodeficiency disorders (e.g. AIDS) autoimmune disorders (e.g.
CC arthritis), neurological disorders (e.g. Alzheimer's disease), metabolic
CC disorders (e.g. phenylketonuria), inflammatory disorders (e.g. asthma),
CC cardiovascular disorders (e.g. atherosclerosis), blood-related disorders
CC (e.g. haemophilia), reproductive disorders (e.g. infertility) and
CC infectious disorders (e.g. Influenza). The polynucleotides of the
CC invention can also be used in gene therapy. AAU22915-AAU23814 represent
CC the novel human enzyme polypeptides of the invention. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX
SQ Sequence 558 AA;

Query Match 48.9%; Score 2595.5; DB 4; Length 558;
Best Local Similarity 94.6%; Pred. No. 3.9e-192;
Matches 510; Conservative 10; Mismatches 16; Indels 3; Gaps 2;

QY 294 ENVAARASPEPELQLRPYQMEVAQPALEGKNIICLTGSGKTRVAVYIAKDHLDKKKA 353
Db 1 ENVAARASPEPELQLRPYQMEVAQPALEGKNIICLTGSGKTRVAVYIAKDHLDKKKA 60
QY 354 SEPGKIVLVNKKVLLVEQLFRKEFPFLKKWYRVIGLSDGTQLKISFPEVVKSCDIIIST 413
Db 61 SEPGKIVLVNKKVLLVEQLFRKEFPFLKKWYRVIGLSDGTQLKISFPEVVKSCDIIIST 120
QY 414 AQILENSLLNLENGEDAGVQLSDPFLIIIDECHHTNKEAVYNNIMRHYLMQKLNRLKK 473
Db 121 AQILENSLLNLENGEDAGVQLSDPFLIIIDECHHTNKEAVYNNIMRHYLMQKLNRLKK 180
QY 474 ENKPVIPPLQILGLTASPGVGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQIQ 533
Db 181 ENKPVIPPLQILGLTASPGVGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQIQ 240
QY 534 EPCKKFAIADATREDPFKEKLEIMTRIQTICYQMSPMSPDFTQPYEQWAIQMEKKAACKG 593
Db 241 EPCKKFAIADATREDPFKEKLEIMTRIQTICYQMSPMSPDFTQPYEQWAIQMEKKAACKG 300
QY 594 NRKERVCAEHLRKYNEALQINDTIRMIDAYTHLETFTYNEEKDKKFAVIEDDSDEGGDEY 653
Db 301 NRKERVCAEHLRKYNEALQINDTIRMIDAYTHLETFTYNEEKDKKFAVIEDDSDEGGDEY 360
QY 654 CDGDEDEDLKKPLKLDLTDRLMTLFFENNMKRLAENPEYENKLTCLRNTIMEQYT 713
Db 361 CDGDEDEDLKKPLKLDLTDRLMTLFFENNMKRLAENPEYENKLTCLRNTIMEQYT 420
QY 714 RTEESARGIIFTKRSAYALSOWITENEFKFAEYGVKAHHLIGAGHSSEFKPMTQNEQKE 773
Db 421 RTEESARGIIFTKRSAYALSOWITENEFKFAEYGVKAHHLIGAGHSSEFKPMTQNEQKE 480
QY 774 VISKFRTGKINLLIATTVAEEGLDIKECNIVIRYGLVTNEIAM-VQARGRARADESTYV 831
Db 481 VISKFRTGKINLLIAPQWKVWILK--NVTLSVMVSSPMKYHGPARGRARADESTYV 537

RESULT 9
ADA54899
ID ADA54899 standard; protein; 468 AA.

XX
AC ADA54899;
XX
DT 20-NOV-2003 (first entry)
XX
DE Human protein, SEQ ID 2467.

XX
KW Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;
KW Gene Therapy; human; secretory protein; membrane proteins; cancer;
KW inflammatory disease; osteoporosis; neurological disease.
XX
OS Homo sapiens.

XX EP1293569-A2.
PN
XX AC
PD
XX 19-MAR-2003.
XX
PF 21-MAR-2002; 2002EP-00006586.
XX
PR 14-SEP-2001; 2001JP-00328381.
PR 24-JAN-2002; 2002US-0350435P.
XX
PA (HELI-) HELIX RES INST.
PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;
XX
DR WPI; 2003-395539/38.
DR N-PSDB; ADA53260.
XX
PT New polynucleotides encoding full-length polypeptides, e.g. secretory
PT and/or membrane proteins, useful for developing medicines for diseases in
PT which the gene is involved, or as target molecules for gene therapy.
XX
PS Claim 14; SEQ ID NO 2467; 205pp; English.
XX
CC The present invention relates to novel human secretory or membrane
CC proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-
CC ADA54071). The coding sequences are useful in the gene therapy of
CC diseases caused by abnormalities of the proteins, e.g. cancer,
CC inflammatory diseases, osteoporosis or neurological disease.
XX
SQ Sequence 468 AA;

Query Match 46.2%; Score 2456; DB 6; Length 468;
Best Local Similarity 99.1%; Pred. No. 2e-181;
Matches 464; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 558 MTRIQTYCOMSPMSDFGTQPYEQWAIQMEKKAAGKGNRERVCAEHLRKYNEALQINDTI 617
Db |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1 MTRIQTYCOMSPMSDFGTQPYEQWAIQMEKKAAGKGNRERVCAEHLRKYNEALQINDTI 60

QY 618 RMIDAYTHLETFYNEEKKKFAVIEDDSDEGGDEYCDGDEDDLKKPLKLDTDREFLM 677
Db |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
61 RMIDAYTHLETFYNEEKKKFAVIEDDSDEGGDEYCDGDEDDLKKPLKLDTDREFLM 120

QY 678 TLFFENNKMRLAENPEYENEKLTKLNTIMEQYTRTESARGIIFTKTRQAYALSQW 737
Db |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
121 TLFFENNKMRLAENPEYENEKLTKLNTIMEQYTRTESARGIIFTKTRQAYALSQW 180

QY 738 ITENEKFAEVGVKAHHLIGAGHSSEFKPMTQNEQKEVISKFRGTGKINLLIATTVAEEGLD 797
Db |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
181 ITENEKFAEVGVKAHHLIGAGHSSEFKPMTQNEQKEVISKFRGTGKINLLIATTVAEEGLD 240

QY 798 IKECNIVIRYGLVTNEIAMVQARGARADESTYLVVHSGSGVIEHETVNDFREKMYKA 857
Db |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
241 IKECNIVIRYGLVTNEIAMVQARGARADESTYLVVHSGSGVIERETVNDFREKMYKA 300

QY 858 IHCQNMPPEYAHKILELQMSIMEKKMKTKRNIKHYKNPSLITFLCKNCSVLACSG 917
Db |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
301 IHCQNMPPEYAHKILELQMSIMEKKMKTKRNIKHYKNPSLITFLCKNCSVLACSG 360

QY 918 EDIHVIEKMHVNMTPEFKELYIVRENKALQKCADYQINGEIIICKGQAWGTMMVHKGL 977
Db |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
361 EDIHVIEKMHVNMTPEFKELYIVRENKTLQKCADYQINGEIIICKGQAWGTMMVHKGL 420

QY 978 DLPCLKIRNFVVFPKNSTKKQYKKWVLPITFPNLDYSECCLFSD 1025
Db |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
421 DLPCLKIRNFVVFPKNSTKKQYKKWVLPITFPNLDYSECCLFSD 468

RESULT 10
AAU23647

ID AAU23647 standard; protein; 417 AA.
XX
AC AAU23647;
XX
DT 18-DEC-2001 (first entry)
XX
DE Novel human enzyme polypeptide #733.
XX
KW Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;
KW ligase; hyperproliferative disorder; immunodeficiency disorder;
KW autoimmune disorder; neurological disorder; metabolic disorder;
KW inflammatory disorder; cardiovascular disorder; reproductive disorder;
KW blood-related disorder; infectious disorder; cytostatic; anti arthritic;
KW nephrotropic; anticoagulant.
XX
OS Homo sapiens.
XX
PN WO200155301-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001239.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.

PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.

PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WFI; 2001-465566/50.
DR N-PSDB; AAS41517.
XX
PT Novel polypeptides and polynucleotides useful for diagnosing, preventing,
PT treating neural, immune system, muscular, reproductive, pulmonary,
PT cardiovascular, renal, proliferative disorders and cancerous diseases.
XX
PS Claim 11; SEQ ID NO 1643; 1180pp; English.
XX
CC The present invention relates to the isolation of novel human enzyme
CC polypeptides, and the cDNA (AAS40785-AAS41684) and genomic sequences
CC encoding them. The enzyme polypeptides of the invention may comprise the
CC functional classes of oxidoreductases, transferases, hydrolases, lyases,
CC isomerases or ligases. The sequences of the invention are useful in the
CC diagnosis, treatment, prevention and/or prognosis of a wide range of
CC disorders including hyperproliferative disorders (e.g. cancer),
CC immunodeficiency disorders (e.g. AIDS), autoimmune disorders (e.g.
CC arthritis), neurological disorders (e.g. Alzheimer's disease), metabolic
CC disorders (e.g. phenylketonuria), inflammatory disorders (e.g. asthma),
CC cardiovascular disorders (e.g. atherosclerosis), blood-related disorders
CC (e.g. haemophilia), reproductive disorders (e.g. infertility) and
CC infectious disorders (e.g. Influenza). The polynucleotides of the
CC invention can also be used in gene therapy. AAU22915-AAU23814 represent
CC the novel human enzyme polypeptides of the invention. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 417 AA;

Query Match 35.0%; Score 1860; DB 4; Length 417;
Best Local Similarity 90.8%; Pred. No. 2.9e-135;
Matches 365; Conservative 13; Mismatches 20; Indels 4; Gaps 3;

QY 442 IDECHHTNKEAVYNNIMRHYLMQKLKNNRLKKENKPVLPQILGLTASPGVGATKQAK 501
Db :|||||
8 VDECHHTNKEAVYNNIMRHYLMQKLKNNRLKKENKPVLPQILGLTASPGVGATKQAK 67
QY 502 AEEHILKLCANLDAFTIKTVKENLDQLKNQIQEPCCKFAIADATREDPFKEKLEIMTRI 561
Db |||||
68 AEEHILKLCANLDAFTIKTVKENLDQLKNQIQEPCCKFAIADATREDPFKEKLEIMTRI 127
QY 562 QTYCQMSPNMDFGTQPYEQWAIQMEKKAAGKGNRKERVCAEHLRKYNEALQINDTIRMID 621
Db |||||
128 QTYCQMSPNMDFGTQPYEQWAIQMEKKAAGKGNRKERVCAEHLRKYNEALQINDTIRMID 187
QY 622 AYTHLETFFYNEEKDKKFAVIEDSDDEGGDDDEYCDGDEDDLKKPLKLDETDRFLMTLFF 681
Db |||||
188 AYTHLETFFYNEEKDKKFAVIEDSDDEGGDDDEYCDGDEDDLKKPLKLDETDRFLMTLFF 247
QY 682 ENNKMLKRLAENPEYENEKLTCLRNTIMEQYTRTEESARGIIFTKTQSAVALSQWITEN 741
Db |||||
248 ENNKMLKRLAENPEYENEKLTCLRNTIMEQYTRTEESARGIIFTKTQSAVALSQWITEN 307
QY 742 EKFAEVGVKAHHLIGAGHSSEFKPMTQNEQKEVISKFRGKINLLIATTVAEEGLDIKEC 801

Db 308 EKFAEVGVKAHHLIGAGHSSEFKPMTQNEQKEVSKFRTGKINLLIAPQWQKVWILK-- 365
QY 802 NIVIRYGLVTNEIAM-VQARGRARADESTYVLVAHSGSGVIE 842
Db 366 NVTLSVMVSSPMKYHGPARGRARADESTYVWLL-QWSGVID 406

RESULT 11
AAM40129
ID AAM40129 standard; protein; 356 AA.
XX
AC AAM40129;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 3274.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US034263.
XX
PR 23-DEC-1999; 99US-00471275.
PR 21-JAN-2000; 2000US-00488725.
PR 25-APR-2000; 2000US-00552317.
PR 20-JUN-2000; 2000US-00598042.
PR 19-JUL-2000; 2000US-00620312.
PR 03-AUG-2000; 2000US-00653450.
PR 14-SEP-2000; 2000US-00662191.
PR 19-OCT-2000; 2000US-00693036.
PR 29-NOV-2000; 2000US-00727344.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
PI Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI; 2001-442253/47.
DR N-PSDB; AAI59285.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders such
as central nervous system injuries.
XX
PS Example 5; SEQ ID NO 3274; 10078pp; English.
XX
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and the
CC encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders. Note: The sequence data for this patent did not form
part of the printed specification
XX

SQ Sequence 356 AA;
Query Match 34.6%; Score 1839; DB 4; Length 356;
Best Local Similarity 99.4%; Pred. NO. 9.6e-134;
Matches 350; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 674 RFLMTLFFENNKKLRLAENPEYENEKLTCLRNTIMEQYTRTEESARGIIFTKTRQSAYA 733
Db 5 RFLMTLFFENNKKLRLAENPEYENEKLTCLRNTIMEQYTRTEESARGIIFTKTRQSAYA 64
QY 734 LSQWITENEKFAEVGVKAHHLIGAGHSSEFKPMTQNEQKEVSKFRTGKINLLIATTVAE 793
Db 65 LSQWITENEKFAEVGVKAHHLIGAGHSSEFKPMTQNEQKEVSKFRTGKINLLIATTVAE 124
QY 794 EGLDIKECNIVIRYGLVTNEIAMVQARGRARADESTYVLVAHSGSGVIEHETVNDFREKM 853
Db 125 EGLDIKECNIVIRYGLVTNEIAMVQARGRARADESTYVLVAHSGSGVIEHETVNDFREKM 184
QY 854 MYKAIHCVQNMKPEEYAHKILELQMQSIMKKMKTNRNIAKHYKKNPSLITFLCKNC SVL 913
Db 185 MYKAIHCVQNMKPEEYAHKILELQMQSIMKKMKTNRNIAKHYKKNPSLITFLCKNC SVL 244
QY 914 ACSGEDIHVIEKMHVNMTPEFKELYIVRENKALQKCADYQINGEIIICKGQAWGTMV 973
Db 245 ACSGEDIHVIEKMHVNMTPEFKELYIVRENKALQKCADYQINGEIIICKGQAWGTMV 304
QY 974 HKGLDLPCLKIRNFVVFKNNSTKKQYKKWVELPITFPNLDYSECCLFSD 1025
Db 305 HKGLDLPCLKIRNFVVFKNNSTKKQYKKWVELPITFPNLDYSECCLFSD 356

RESULT 12
AAM41915
ID AAM41915 standard; protein; 348 AA.
XX
AC AAM41915;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 6846.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US034263.
XX
PR 23-DEC-1999; 99US-00471275.
PR 21-JAN-2000; 2000US-00488725.
PR 25-APR-2000; 2000US-00552317.
PR 20-JUN-2000; 2000US-00598042.
PR 19-JUL-2000; 2000US-00620312.
PR 03-AUG-2000; 2000US-00653450.
PR 14-SEP-2000; 2000US-00662191.
PR 19-OCT-2000; 2000US-00693036.
PR 29-NOV-2000; 2000US-00727344.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
PI Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI; 2001-442253/47.

QY 800 ECNIVIRYGLVTNEIAMVQARGRADESTYVLVAHSGSGVIEHETVNDFREKMYKAH 859
Db 123 ECNIVIRYGLVTNEIAMVQARGRADESTYVLVAHSGSGVIEHETVNDFREKMYKAH 182
QY 860 CVQNMKPEEYAHKILELOMQSIMEKMKTKRNIKHYKNPDLITFLCKNCSVLACSGED 919
Db 183 CVQNMKPEEYAHKILELOMQSIMEKMKTKRNIKHYKNPDLITFLCKNCSVLACSGED 242
QY 920 IHVIEKMHVNMTPEFKELYIVRENKALQKKCADYQINGEICKCGQAWGTMVHKGLDL 979
Db 243 IHVIEKMHVNMTPEFKELYIVRENKTLQKKCADYQINGEICKCGQAWGTMVHKGLDL 302
QY 980 PCLKIRNFVVVFKNSTKKQYKKWVELPITFPNLDYSECCLFSD 1025
Db 303 PCLKIRNFVVVFKNSTKKQYKKWVELPITFPNLDYSECCLFSD 348

RESULT 14

ABP69672
ID ABP69672 standard; protein; 678 AA.

XX ABP69672;

AC
XX 20-JAN-2003 (first entry)

DT Human polypeptide SEQ ID NO 1719.

XX Human; genome mapping; gene therapy; food supplement; virus; fungus;
KW cell-proliferative disorder; neurodegenerative disease; bacterial;
KW Parkinson's disease; Alzheimer's disease; autoimmune disease;
KW multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;
KW arthritis; cytostatic; immunomodulator; nootropic; neuroprotective;
KW antiparkinsonian; antidiabetic; immunosuppressive; dermatological;
KW haemostatic; vulnery; fungicide; antibacterial; virucide; protozoacide;
KW antiarthritic.

XX Homo sapiens.

XX WO200270539-A2.

PN 12-SEP-2002.

XX 05-MAR-2002; 2002WO-US005095.

XX 05-MAR-2001; 2001US-00799451.

XX (HYSE-) HYSEQ INC.

PI Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;
PI Wehrman T, Wang J, Wang D, Drmanac RT;

XX WPI; 2002-759812/82.

DR N-PSDB; ABZ11889.

XX New polynucleotides comprising sequences assembled from expressed
PT sequence tags (ESTs), useful for treating cell-proliferative,
PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or platelet
PT or coagulation disorders.

XX Claim 9; SEQ ID NO 1719; 1012pp + Sequence Listing; English.

XX The invention relates to an isolated polynucleotide (I) comprising a
CC nucleotide sequence selected from any of 948 sequences (ABZ11119-
CC ABZ12066) or their mature protein coding portion, active domain coding
CC protein or complementary sequences. The polynucleotides are useful for
CC identifying expressed genes or for physical mapping of human genome. The
CC encoded polypeptides (ABP68902-ABP69849) are useful as molecular weight
CC markers, as a food supplement, for generating antibodies, in medical
CC imaging, screening and diagnostic assays and for treating cell-
CC proliferative disorders (cancer), neurodegenerative diseases (Parkinson's
CC or Alzheimer's disease), autoimmune diseases (multiple sclerosis,
CC diabetes, lupus) genetic disorders, myeloid or lymphoid disorders,

CC platelet or coagulation disorders, wound, burns, incision, ulcers, liver
CC or lung fibrosis, infections (bacterial, viral, fungal, parasitic),
CC arthritis, etc. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 678 AA;

Query Match 25.5%; Score 1352; DB 5; Length 678;
Best Local Similarity 40.7%; Pred. No. 1.6e-95;
Matches 292; Conservative 136; Mismatches 237; Indels 52; Gaps 12;

QY 306 LQLRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKDHLDKKKKASEPGKVIVLVNK 365
Db 1 MELRSYQWEVIMPALLEGKNIILWPTGAGKTRAAAYVAKRHLE---TVDGAKVVVLNR 56
QY 366 VLLVEQLFRKEFQFLKKWYRVIGLSGDTQLKISFPEVVKSCDIIISTAQILENSLLNLE 425
Db 57 VHLVTQ-HGEFFRRMLDGRWTVTTLSGDMGPRAGFGLARCHDLLICTAELLQMALTSPE 115
QY 426 NGEDAGVQLSDFSLIIDECHHTNKEAVYNNIMRHYLMQKLKNNRLKKNKVPVPLPQIL 485
Db 116 --EEHVELTVFSLIVVDECHHTHKDTVYNNVIMSQYLELKLQRAQ-----PLPQVL 164
QY 486 GLTASPGVGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQIQEPCKKFAIADAT 545
Db 165 GLTASPGTGASKLDGAINHVQLCANLDTWCIMSPQNCPPQLQHSQQPCKQYNLCHRR 224
QY 546 REDPFKEKLEIMTRIQTVCQSPMS-DFGTQPYEQWAIQMEKKAAXKGNRKERVCAEHL 604
Db 225 SQDPFGDLLKKLMDQIHDHLEMPELSRKFGTQYEQVQVVKLSEAAALAGLQEQRVYALHL 284
QY 605 RKYNEALQINDTIRMIDAYTHLETIFYNEEKDKKFAVIEDDSDEGGDDCYDGEDDDLK 664
Db 285 RRYNDALLIHDTVRAVDALAALQDFYHREHVTQTIL-----C----- 322
QY 665 KPLKLDDETRFLMTLFFENNMKRLKLAENPEYENEKLTCLRNTIMEQYTRTEESARGIIF 724
Db 323 -----AERRLLALFDDRKNELAHLATGCP-ENPKLEMLEKILQRFSS-SSNSPRGIIF 373
QY 725 TKTRQSAYALSQWITENEKPAEVGVKAHLIGAGHSSEFKPMTQNEQKEVISKFTGKIN 784
Db 374 TRTRQSAHSLLLWLOQQOGLQTVDIRAQLLIGAGNSSQSTHMTQRDQOEVIQKFDGTLN 433
QY 785 LLIATTVAEEGLDIKECNIVIRYGLVTNEIAMVQARGRADESTYVLVAHSGSGVIEHE 844
Db 434 LLVATSVAAEEGLDIPHCNVVVRVRYGLLTNEISMVQARGRARADQSVYAFVATEGSRELKRE 493
QY 845 TVNDFREKMYKAHCVQNMKPEEYAHKILELQMQSIMEKMKTKRNIKHYKNPDLIT 904
Db 494 LINEALETLMEQAVAAVQKMDQAEYQAKIRDLQQAALTKRAAQAAQRENQRQFPVEHVQ 553
QY 905 FLCKNCSVLACSGEDIHVEIKMHVNMTPEFKELY-IVRENKALQKKCADYQINGEICK 963
Db 554 LLCINCMVAVGHGSLDKVGEHTHHVNVNPNFNSYNNVSRDPVIVNKVFKDWKPGGVISCR 613
QY 964 -CGQAWGTMVHKGLDLPCLKIRNFVVVFKNSTKKQYKKWVELPITFPNLDYSECC 1019
Db 614 NCGEVWGLQMIYKSVKLPVLKVR--MLLETPQGRIOAKKWSRVFSPVDFDFLQHC 668

RESULT 15

AAB93708
ID AAB93708 standard; protein; 678 AA.

XX AAB93708;

AC AAB93708;

XX 26-JUN-2001 (first entry)

DT Human protein sequence SEQ ID NO:13299.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 26, 2005, 20:07:35 ; Search time 42 Seconds
(without alignments)
1618.475 Million cell updates/sec

Title: US-09-515-363C-2
Perfect score: 5311
Sequence: 1 MSGYSTDENFRYLISCFRA.....LPITFPNLDYSECCLFSDDED 1025

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	406	7.6	1909	4	US-09-590-968B-2
2	286	5.4	503	4	US-09-408-020-66
3	284	5.3	502	4	US-09-408-020-34
4	217.5	4.1	2662	4	US-09-595-684B-31
5	217.5	4.1	2663	4	US-09-538-092-1252
6	202	3.8	3878	4	US-09-914-259-11
7	194.5	3.7	2482	1	US-08-328-254-6
8	193.5	3.6	661	4	US-09-107-532A-3677
9	191.5	3.6	3248	1	US-08-353-700-1
10	191.5	3.6	3248	5	PCT-US95-16216-1
11	189.5	3.5	3210	4	US-09-538-092-1154
12	188.5	3.5	1388	2	US-08-685-576-4
13	188.5	3.5	1388	4	US-09-976-594-296
14	187.5	3.5	2125	4	US-09-919-172-29
15	186.5	3.5	1786	3	US-08-973-462-8
16	184	3.5	317	4	US-09-248-796A-15035
17	182.5	3.4	2954	4	US-09-150-867-1
18	180.5	3.4	78	4	US-09-590-968B-5
19	180.5	3.4	976	4	US-09-538-092-1339
20	180	3.4	2704	4	US-09-538-092-1260
21	179	3.4	662	4	US-09-583-110-4571
22	179	3.4	666	3	US-09-134-001C-5465
23	175	3.3	81	4	US-09-590-968B-9
24	174	3.3	1388	2	US-08-685-576-1
25	173	3.3	1211	3	US-09-134-001C-4820
26	173	3.3	2349	4	US-09-538-092-914
27	171.5	3.2	976	3	US-09-104-324B-4

28	171.5	3.2	2285	3	US-09-308-375-2	Sequence 2, Appli
29	171	3.2	1401	4	US-09-750-590-2	Sequence 2, Appli
30	170	3.2	905	4	US-09-248-796A-16333	Sequence 16333, A
31	167.5	3.2	1132	4	US-09-248-796A-15026	Sequence 15026, A
32	167.5	3.2	1219	4	US-09-107-532A-6020	Sequence 6020, Ap
33	167	3.1	677	4	US-09-328-352-4365	Sequence 4365, Ap
34	167	3.1	3696	3	US-09-134-001C-5080	Sequence 5080, Ap
35	166.5	3.1	1886	3	US-08-938-105-3	Sequence 3, Appli
36	166	3.1	810	4	US-09-248-796A-20281	Sequence 20281, A
37	166	3.1	1937	4	US-09-538-092-918	Sequence 918, App
38	164	3.1	1354	3	US-08-685-871-2	Sequence 2, Appli
39	163	3.1	1939	3	US-09-310-187A-1	Sequence 1, Appli
40	162	3.1	1939	4	US-09-538-092-917	Sequence 917, App
41	159.5	3.0	956	3	US-09-134-001C-4452	Sequence 4452, Ap
42	159	3.0	442	4	US-09-489-039A-11770	Sequence 11770, A
43	158	3.0	781	2	US-08-675-631-3	Sequence 3, Appli
44	158	3.0	781	4	US-08-359-316A-3	Sequence 3, Appli
45	158	3.0	781	4	US-09-248-776-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1

US-09-590-968B-2
; Sequence 2, Application US/09590968B
; Patent No. 6737561
; GENERAL INFORMATION:
; APPLICANT: Ray, Animesh
; APPLICANT: Golden, Teresa Ann
; TITLE OF INVENTION: GENE ENCODING SHORT INTEGRUMENTS AND USES THEREOF
; FILE REFERENCE: 176/60581
; CURRENT APPLICATION NUMBER: US/09/590,968B
; CURRENT FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/138,316
; PRIOR FILING DATE: 1999-06-09
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1909
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-590-968B-2

Query Match		7.6%;	Score 406;	DB 4;	Length 1909;
Best local Similarity		23.0%;	Pred. No. 3.4e-25;		
Matches 199;		Conservative 131;	Mismatches 271;	Indels 264;	Gaps 40;
QY	135	LDKCM-----EELLTIEDNRIRIAAENNGNESGVRELLKRIVQKENWFSAFNLVLRQT	188		
Db	18	LDACEDISCDLIDLVSEFDPSSVAVNEST-DENG-----INDFGGIDHILDSI	67		
QY	189	GNNELVQELTGCSESNAETENLSQVDGPQVEEQLLSTTVQPN-LEKEVWGMENNSSSES	247		
Db	68	KNG---GGLPNNGVSDTNSQINEVTVT--PQV---IAKETVKEGLQKNGGKRDEFSKEE	119		
QY	248	SFAD-----SSVSESDTSLAEGSVSCLDE-----SLGHN-----	277		
Db	120	GDKDRKRARVCSYQSERSNLGRGHVNNREGDRFMNRKRTRNWDAGNNKKKRECNMYR	179		
QY	278	-----SNMGSD-----SGTMGSDSD-----EENVAAAS	301		
Db	180	RDGRDREVRGYWERDKVGSNELVYRSGTWEADHERDVKKVSGGNRECDVKAEN---KSK	236		
QY	302	PEP-----ELQLRPYQMEVAQPALEGKNIICLPTSGCKTRVAVY-----IAKDLDKKK	351		
Db	237	PEERKEKVVEEQARRYQLDLEQA-KAKNTIAFLETGAGKTLIAILLIKSVHKDLMSONR	295		
QY	352	KASEPGKIVILVKNVLLVEQ---LFRKEF-----QPF--LKKWYRVIGLSGDT	394		
Db	296	KMLS-----VFLVPKVPLVYQQAQAEVIRNQTCFQVGHYCGEMGQDFWDSRRWQREF-----	345		
QY	395	QLKISFPEVVKSCDIIISTAQILENSLLNLENGEDAGVQLSDFSLSIIIDECHHTNKEAVY	454		

Db 346 -----ESKQVLVMTAQILNLL-----RHSIIRMETIDLLILDECHHAVKHPY 389
QY 455 NNIMRHYLMOKLKNRLKKNKPVIPLPQILGLTASP-GVGGATKQAKAEHHILKLCANL 513
Db 390 SLVMSEFYHTTPKDKR-----PAIFGMTASPVNLKGVSSQVDCAIKIRNLETKL 438
QY 514 DAFTIKTVKENLDQLKNQIQEPCKKFAIAD-ATREDPFKEKLEIMTRIQTQYCMSPMSD 572
Db 439 DS-TVCTIKDR-KELEKHVPMPSEIVVEYDKAATWSLHETIKQIAAVEEAAQASSRK- 495
QY 573 FGTQPYEQWAIQMEKKAAGKGNRKERVCAEHLRKYNEALQINDTIRMID-----A 622
Db 496 -----SKWQFMGARDAGAKDELROVYGVVSERTESDGAANLIHKLRAINITYTLAELGOWCA 549
QY 623 YTHLETF-----YNEEKDKKF-----AVIE----- 642
Db 550 YKVGQSFLSALQSDERVNFQVDVKFQESYLSEVSLLOCELLEGAAAEKVAAEVGKPPENG 609
QY 643 ---DDSDEG-GDDEYCDGDEDEDLKKPLKLDETDRFLMTLFFENNKMLKRLAENPEYE 697
Db 610 NAHDEMEGELPDDPVVSGGEHVDEVIGAAVAD-----GKVTPKV----- 649
QY 698 NEKLTCLRNTIMEQYTRTEESARGIIFTKTRQSAYALSQWITENEKFAEVG----VKAHH 753
Db 650 -QSLIKL----LLKYQHTAD-FRAIVFVERVVAALVP-----KVFAELPSLSFIRCAS 697
QY 754 LIGAGHSSFPKPMWTQNEQKEVISKFTGKINLLIATTVAEEGLDIKECNIVIRYGLVTNE 813
Db 698 MIGHNNSQEMK---SSQMQDTISKFRDGHVTLVATSVAAEEGLDIRQCNVVMRFDLAKTV 754
QY 814 IAMVQARGARADESTYVLVAHSGS 838
Db 755 LAYIQSRGRARKPGSDYILMVERGN 779

RESULT 2
US-09-408-020-66
; Sequence 66, Application US/09408020
; Patent No. 6632937
; GENERAL INFORMATION:
; APPLICANT: Swanson, Ronald V.
; APPLICANT: Feldman, Robert A.
; APPLICANT: Schleper, Christa
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM
; FILE REFERENCE: DCOIP.002A
; CURRENT APPLICATION NUMBER: US/09/408,020
; CURRENT FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: 60/102,294
; PRIOR FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 66
; LENGTH: 503
; TYPE: PRT
; ORGANISM: Cenarchaeum symbiosum.
US-09-408-020-66

Query Match 5.4%; Score 286; DB 4; Length 503;
Best Local Similarity 24.7%; Pred. No. 8.5e-16;
Matches 136; Conservative 91; Mismatches 204; Indels 120; Gaps 24;
QY 303 EP-ELQLRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKDHLDKKKKASEPGKVIV 361
Db 12 EPGAVERRDYQVGLAEQAIR-ENCIVLPTGLGKTAVALQVISHYLDDEGRGALFLAPTRV 70
QY 362 LVNKKVLLVEQLFRKEFPFLKKWYRVIGLSGDTQL--KISFPEVVKSC--DIIISTAQIL 417
Db 71 LVN-----QHRQFLG---RALTISDITLVTGEDTVPRRKKAWGGSVICATPEIT 116
QY 418 ENSLLNLENGEDAGVQLSDFSLLIIDECHHTNKEAVYNNIMRHYLMQKLNRLKKNKP 477
Db 117 RNDIAR-----GMVPLEQFGLVVFDEAHRAVGDYAYSIAA-----RAVGENS- 158

QY 478 VIPLPQILGLTASPGVGGATKQAKAEHHILKLCANLDAFTIKTVKENLDQLKNQIQEPCK 537
Db 159 -----RMIGMTATL-----PSEREKADE---IMGTLLSKSIAORTEDDDPDVKPYVQETET 205
QY 538 KFAIADATREDPFKEKLEIMTRIQTQYCMSPMS-DFGTQPYEQWAIQMEKKAAGKGNRK 596
Db 206 EWIKVELPPMEMKEIQKLLK-MALDERYAALKRCGYDLGSNRSLSALLRL-RMVVLSGNRR 263
QY 597 ERVCAEHLRKYNEALQINDTIRMIDAY--THLETIFYNEEKDKKFAVIEDDSDEGGDDEYC 654
Db 264 A-----AKPLFTAIRITYALNIFEAHGVTVPFLKFCERTVKKKGAGVAE----- 306
QY 655 DGDEDEDDLKPLKLDETDRFLMTLFFENNKMLKRLAENPEYENKLTCLRNTIMEQYTR 714
Db 307 -----LFEEDR-----NFTGAMARAKAAQAAGMEHPKIPKL----- 337
QY 715 TERSARG-----IIFTKTRQSAYALSQWITENEKFAEVGVKAHHLIG-AGHSSEFKPMTQ 768
Db 338 -EEAVRGAKGKALVFTSYRDSVDLI-----HSLQAAGINSILIGKAGE----KGLKQ 386
QY 769 NEQKEVISKFTGKINLLIATTVAEEGLDIKECNIVIRYGLVTNEIAMVQARGR-ARADE 827
Db 387 KKQVETVAKFRDGGYDVLVSTRVGEGLDISEVNLVVFYDNVPSSIRYVQRRGRTGRKDA 446
QY 828 STYVLVAHSGS 838
Db 447 GKLVVLMAGKT 457

RESULT 3
US-09-408-020-34
; Sequence 34, Application US/09408020
; Patent No. 6632937
; GENERAL INFORMATION:
; APPLICANT: Swanson, Ronald V.
; APPLICANT: Feldman, Robert A.
; APPLICANT: Schleper, Christa
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM
; FILE REFERENCE: DCOIP.002A
; CURRENT APPLICATION NUMBER: US/09/408,020
; CURRENT FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: 60/102,294
; PRIOR FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 34
; LENGTH: 502
; TYPE: PRT
; ORGANISM: Cenarchaeum symbiosum
US-09-408-020-34
Query Match 5.3%; Score 284; DB 4; Length 502;
Best Local Similarity 24.3%; Pred. No. 1.3e-15;
Matches 133; Conservative 94; Mismatches 207; Indels 114; Gaps 23;
QY 303 EP-ELQLRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKDHLDKKKKASEPGKVIV 361
Db 12 EPGAVERRDYQVGLAEQAIR-ENCIVLPTGLGKTAVALQVIAHYLDDEGRGALFLAPTRV 70
QY 362 LVNKKVLLVEQLFRKEFPFLKKWYRVIGLSGDTQL--KISFPEVVKSC--DIIISTAQIL 417
Db 71 LVN-----QHRQFLG---RALTISDITLVTGEDTIPRRKKAWGGSVICATPEIA 116
QY 418 ENSLLNLENGEDAGVQLSDFSLLIIDECHHTNKEAVYNNIMRHYLMQKLNRLKKNKP 477
Db 117 RNDI-----ERGLVPLEQFGLVVFDEAHRAVGDYAYSIIAR-----A 153
QY 478 VIPLPQILGLTASPGVGGATKQAKAEHHILKLCANLDAFTIKTVKENLDQLKNQIQEPCK 537
Db 154 VGDNSRMVGMGTATL-----PSEREKADE---IMGTLLSRSTIAORTEDDDPDVKPYVQETAT 205
QY 538 KFAIADATREDPFKEKLEIMTRIQTQYCMSPMS-DFGTQPYEQWAIQMEKKAAGKGNRK 596

Db 206 EWIKVDLPPEMKEIQRLLLKALD-ERYSSLKRCGYDLGSNRSLALLRL-RMVVLGNRR 263
QY 597 ERVCAEHLRKYNEALQINDTIRMIDAY--THLETFFYNEEKDKKFAVIEDDSDEGGDEYC 654
Db 264 A-----AKPLFTAIRITYALNIFEAHGVTFFLKFCERTSKKKGVGVAE----- 306
QY 655 DGDEDEDLKPLKLDLDRFLMTLFFENNKMRLAENPEYENEKLTCLRNTIMEQYTR 714
Db 307 -----LFEQDR-----NFTGAIRAKAAQAAGMEHPKIPKLEDAV----- 341
QY 715 TEESARG--IIFTKTRQSAYALSOWITENEKFAEVGVKAHHLIG-AGHSSEFKPMTQNEQ 771
Db 342 --RGARGKALVFTSYRDSVDLI-----HSRLKAAGINSGLIGKAGE----KGLKQRKQ 389
QY 772 KEVSKFRGKINLLIATTVAEEGLDIKECNIVIRYGLVTNEIAMVQARGR-ARADESTY 830
Db 390 VETVAKFRDGGYDVLVSTRVGEGLDISEVNLVIFYDNPSSIRYVQRRGRTGRKDAGRL 449
QY 831 VLVAHSGS 838
Db 450 IVLMAKGT 457
RESULT 4
US-09-595-684B-31
; Sequence 31, Application US/095955684B
; Patent No. 6544766
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Ohashi, Cara
; APPLICANT: Sakowicz, Roman
; APPLICANT: Vaisberg, Eugeni
; APPLICANT: Wood, Kenneth
; APPLICANT: Yu, Ming
; TITLE OF INVENTION: Human kinesins and methods of producing
; TITLE OF INVENTION: and purifying human kinesins
; FILE REFERENCE: cytop036
; CURRENT APPLICATION NUMBER: US/09/595,684B
; CURRENT FILING DATE: 2002-06-24
; PRIOR APPLICATION NUMBER: 09/295,612
; PRIOR FILING DATE: 2000-04-20
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 31
; LENGTH: 2662
; TYPE: PRT
; ORGANISM: Human
US-09-595-684B-31
Query Match 4.1%; Score 217.5; DB 4; Length 2662;
Best Local Similarity 18.6%; Pred. No. 9.7e-09;
Matches 209; Conservative 185; Mismatches 421; Indels 307; Gaps 47;
QY 9 ENFRYLISCFARVKMYIQVEPVLDTLFLPAEVKEQIQRTVATSGNMQAVELLSTLEK 68
Db 1039 EQQRKIFSLIQEKVELQOMLESVIAEKEQLKTDLKENIEMTIE--NQEELRLIGDELKK 1095
QY 69 GVWHLGWTRFV----EALRRTGSPLAARYMNPDLTDLSPSPFENAHDEYLQLNLLQP 123
Db 1096 -----QQEIVAQEKHAIKKEGELSRTCRLAEVVEEKLKEKSQQLOEKQQQLLNQVEE 1148
QY 124 TLVDKLLVRDV--LDKMEEEELLTIE--DRNRIAAEN-NGNESGVRELLK-RIVQKE-- 175
Db 1149 MSEMOKKINEIENLKNELKNKELTLEHETERLELAQKLNENYEEVKSITKERKVLKELQ 1208
QY 176 -----NWFSAPLNLVLRQTG-----NNELVQELTGDSCSESABIEENL 212
Db 1209 KSFETERDHLRGYIREIATGLQTKBELKIAHILHKEHQETIDELRRS-VSEKTAQLINT 1267
QY 213 SQVDGPQV-----EEQLLSTTVQPNLEKEVWGMENNSSESSFADSSVVVSESPTS 261
Db 1268 QDLEKSHTKLQBEIPVLVHEEQELL-----PNVKVSVTQETMNELELLTEQSTTKDS-TT 1321

QY 262 LAESVSCLDESGLHNSNMGSDSGTMGSDSDEENVAARASPEPELQLRPYQMEVAQPALE 321
Db 1322 LARIEM-----ERLRLNEKF-----QESQEEISLTKERDNLKTIKEALEVXHDQL- 1367
QY 322 GKNIIICLPTGSGKTRVAVIAK-----DHLDKKKKASEPGKVIVLVNKKVLLVEQL 372
Db 1368 -----KEHIRETLAKIQESQSQEQSLNNKEKDNETTKI-----VSEMEQ- 1407
QY 373 FRKEPQFLKKWYRVIGSLGTQLKISFPE---VVKSCDIIISTAQILENSLLNLENG-- 427
Db 1408 FKPKDSALLRIEIEMLGLS--KRLQESHDEMSVAKEKDDLQRLQEVLQSESDQKENIK 1465
QY 428 EDAGVOLSDFSLIIDECHHTNKEAVYNNIMRHYLMQKLKNRNLKKNKPKVPLPQILGL 487
Db 1466 EIVAKHLETEEELKVAHCCLKEQEBETIN-----ELRVNLSEKETE----- 1505
QY 488 TASPVGATKQAKA-----BEHILKLCANLDAFTIKTVKENLDQLKNQIQEPCKKF 539
Db 1506 -----ISTIQQLEAINDKLQNKIQEIYEKEEQNLNKQISEVQENVNELK-QFKEHRK-- 1557
QY 540 AIADATREDPFKEKLLLEIMTRIQTQYCO--MSPMSDFGTQPYEQWAIQMEKKAACKGNR-- 595
Db 1558 --AKDSALQSIESKMLELTNRLQESQEEIQIMIKEEMKRVQEQALQIERDQLKENTKEI 1615
QY 596 -----KERVCAEHLRKYNEALQIN-----DTIRMID-AYT 624
Db 1616 VAKMKEQEKEYQFLKMTAVNETQEKMCIEIHLKEQFETQKLNLENIETENIRLTQILHE 1675
QY 625 HLETFFYNEEKDKKFAVIEDSDSGDDEYCDGDEDEDLKKPLKLD----- 670
Db 1676 NLEEMRSVTKER-----DDLRSVEETLKVERDQKLENLRETITRDLEKQEELKIVMHML 1729
QY 671 ----ETDRFLMTLFFENNKMRLAENPEYENELK-----TKLRNTIMEQYTRTE 716
Db 1730 KEHQETIDKLRGIVSEKTNESNMQKDLHSNDALKAQDLKIQEBELRIAHHLKEQETI 1789
QY 717 ESARGIIFTKTRQSAYALSQWITEN-----EKFAEVGVKAHHLIGAGHS-----SEFK 764
Db 1790 DKLRGIVSEKTDKLSNMQKDLSENSNAKLOEKIQELKANEHQLITLKKDVNETQKKVSEME 1849
QY 765 PMTQ--NEQKEVISKFRGTGKINLLIATTVAEEGLDIKECNIVIRYGLVTNEIAMVQARGR 822
Db 1850 QLKQIKDQSLTSLKLEIENLNL--AQELHENLEEMK-----SYMKERDN 1892
QY 823 ARADESTYVLVAHSGSGVI-----EH-ETVNDFREKMYKAIH 859
Db 1893 LRRVEETLKLERDQLKESLQETKARDLEIQQELKTARMLSKHEKETVDKLRKISEKTIQ 1952
QY 860 CVQNMK-----PEEYAHKILELQMQSIM-----EKKMKTKRNIKHYKNPSLI 903
Db 1953 ISDIQKDLDKSKDELQKKIQELQKKELQLLLRVKEDVNMSHKKINEMEQKKQFEPN---- 2008
QY 904 TFLCKNCVLCAGEDIHVIEKMHVNMTPFEFKELIYVRENK 945
Db 2009 -YLCK-----CEMDNFQLTKLHE-----SLSEIRIVAKER 2038

RESULT 5

US-09-538-092-1252
; Sequence 1252, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01

; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuraPatSeqFormatter Version 0.9
; SEQ ID NO 1252
; LENGTH: 2663
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number Q02224
US-09-538-092-1252

Query Match 4.1%; Score 217.5; DB 4; Length 2663;
Best Local Similarity 18.6%; Pred. No. 9.7e-09;
Matches 209; Conservative 185; Mismatches 421; Indels 307; Gaps 47;

QY 9 ENFRYLISCFRARVMYIQVEPVLDYLTFLPAEVEKEIQRTVATSGNMQAVELLSTLEK 68
Db 1040 EQQRKIFSLIQKNELQOMLESVIAEKEQLKTDLKENIEMTIE---NQEELRLLDGLKK 1096
QY 69 GVWHLGWTREFV-----EALRRGTSPLAARYMNPETLPLSPSPFENAHDEYQLNLNLLQP 123
Db 1097 -----QQEIVAQEKNAHKAKEGELSKRTCDRLAEVEEKLKESQQLQEKQQQLLNQVEE 1149
QY 124 TLVDKLLVRDV--LDKCMEEELLTIE--DRNRIAAAEN-NGNESGVRELLK-RIVQKE-- 175
Db 1150 MSEMQKKINEIENLKNELKNKELTLEHMETERLELAQKLNENYEEVKSITKERVKLQ 1209
QY 176 -----NWFSAPLNVLRQTG-----NNELVQELTSGDCSESNAEINL 212
Db 1210 KSFETERDLRGVIREIATGLQTKBELKIAHHLKEHQETIDELRRS-VSEKTAQIINT 1268
QY 213 SQVDGPQV-----EEQLLSTTVQPNLEKEVWGMNNSSESFADSVVSESDTS 261
Db 1269 QDLEKSHTKLQEEIPVLHEEQELL-----PNVKKVSEIQTMTNELELLTEQSTTKDS-TT 1322
QY 262 LAEGSVSCLDESIGHNSNMGSDSGTMGSDSDEENVAARASPEPELQRLRPYQMEVAQPALE 321
Db 1323 LARIEM-----ERLRLNEKF-----QESQEEIKSLTKERDNLKTIKEALEVKHDQL- 1368
QY 322 GKNIIICLPTSGKTRVAVYIAK-----DHLDDKKKASEPGKVIVLVNKNVLLVEQL 372
Db 1369 -----KEHIRETLAKIQESQSQEQSLNMBKDNETTKI-----VSEMEQ- 1408
QY 373 FRKEFQPLKWKYRVIGLSDGTQLKISFPE---VVKSCDIIISTAQILENSILLNLENG-- 427
Db 1409 FKPKDSALLRIEIEMLGLS--KRLQESHDEMKSVAKBKDDLQRLQEVLOSQSDQKENIK 1466
QY 428 EDAGVQLSDFSLLIIDECHHTNKEAVYNNIMRHYLMQKLKNRLKKNKPIPLPQILGL 487
Db 1467 EIVAKHLEETEELKVAHCCLKEQETIN-----ELRVNLSEKETE----- 1506
QY 488 TASPVGGAATQAKA-----EEHLKLCANLDAFTIKTVKENLDQLKNQIQEPCKKF 539
Db 1507 -----ISTIQKQLEAINDKLQNKIQEIYEKEEQNLNIKQISEVQENVNELK-QFKEHRK-- 1558
QY 540 AIADATREDPFKEKLLLEIMTRIQTQYQ--MSPMSDFGTQPYEQWAIQMEKKAACKGNR-- 595
Db 1559 --AKDSALQSIESKMLELTNRLQESQEEIQIMIKEEMKRVQEQALQIERDQLKENTKEI 1616
QY 596 -----KERVCAEHLRKVNEALQIN-----DTIRMID-AYT 624
Db 1617 VAKMKESQEKYQFLKMTAVNETQKMCIEIHLKEQFETQKLNLENITENIRLTQILHE 1676
QY 625 HLETFYNEEKDKKFAVIEDSDDEGGDDEYCDGDEDEDDKKPLKD----- 670
Db 1677 NLEEMRSVTKER-----DDLRSVEETLKVVERDQKENLRETITRDLEKQEELKIVHMH 1730
QY 671 -----ETDRFLMTLFFENNKMMLKRLAENPEYENKEL-----TKLRNTIMEQYTRTE 716
Db 1731 KEHQETIDKLRGIVSEKTNESINMQKOLEHSNDALKAQDLKIQEELRIAHMHLKEQETI 1790
QY 717 ESARGIIFTKTRQSAVALSQWITEN----EKFAEVGVKAHHLIGAGHS-----SEFK 764

RESULT 6
US-09-914-259-11
; Sequence 11, Application US/09914259
; Patent No. 6495336
; GENERAL INFORMATION:
; APPLICANT: Makowski, Lee
; APPLICANT: Hyman, Paul
; APPLICANT: Williams, Mark
; TITLE OF INVENTION: STAGED ASSEMBLY OF NANOSTRUCTURES
; FILE REFERENCE: 8471-010-999
; CURRENT APPLICATION NUMBER: US/09/914,259
; CURRENT FILING DATE: 2000-11-21
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 3878
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-914-259-11

Query Match 3.8%; Score 202; DB 4; Length 3878;
Best Local Similarity 18.8%; Pred. No. 3.8e-07;
Matches 238; Conservative 179; Mismatches 424; Indels 428; Gaps 55;

QY 5 YSTD-----ENFRYLISCFRARVMYIQVEPVLDYLTFLPAEVEKEIQRTVATSGNM 56
Db 1136 YSTHVDQVREYNENEKDKALCSLKEELIFAQEEKIKELQKHQLEL--QTMKTQETGDEG 1193
QY 57 QAVELLSTLEKGVWHLGWTREFVEALRRRTGSPLAARYMNPETLPLSPSPFENAHDEYL 115
Db 1194 KPLHLLIGKLQKAV-----SECSYFLQTLCSVLGEYTPALKCEVNAEDKENSGLD--- 1244
QY 116 QLLNLLQPTLVD-KLLVRD-----VLDKCMEE--ELLTIEDR-NRIAAENNGN--ES 162
Db 1245 YISENEDPELDYRYEVDQFQENMHTLLNKVTEYNKLLVLQTRLSKIWGQQTDGKMLEF 1304
QY 163 GVRELLKRIQKENWFSAFNLVLRQTGNNE-----LVQELTSGDCSESNAEIN 211
Db 1305 GEENLPKEETE-----FLSIHSQMTNLEDIDVNHKSKLSSLODLEKTKLEEQVQELES 1357
QY 212 LSQVDGPGQVEEQLLSTTVQPNLEKEVWGMENN---SSBSSF-----ADSSVVSSEDTSLA 263
Db 1358 LIS-----SLQQQLKET--EQNYEABIHCLQKRLQAVSESTVPPSLPVDVSVVITESDAQRT 1411
QY 264 EGSVSCLDSESLGHNSNMGSDSGTMGSDSDEENVAARASPEPELQRLRPYQMEVAQPALEGK 323
Db 1412 MYPGSCVKKNIDGTIEFSGFEGV----KEETNIV-----KLEKQYQEQLEE---EVA 1457
QY 324 NIICLPL-----TGSGKTRVAVYIAKDHLDKKKASEPGKVIVLVNKN 365
Db 1458 KVIIVMSIAFAQQTELRSRISGCKENTASSKQAHAVCQEQEHYFNEMKLSQDQ---IGFQT 1514
QY 366 VLLVEQLFRKEFPFLKKWYRVIGLSG-DTQLKISFPFVV---KSC----- 407

Db 1515 FETVDVKFEKFKPLSKE---LGEHGKILLSNSDPHDIPESKDCVLTISEEMFSKDKT 1570
QY 408 -----DIIISTAQILENSLLNLENGEDAGVQL-----SD 436
Db 1571 FIVRQSIHDEISVSSMDASRQLMLNEEQLEDMRQELVRQYQEHQATQRSSIDNENLVSE 1630
QY 437 FSLIIDE-----CHHTNKEAVYNN----- 456
Db 1631 RERVILLEALEALKQLSLAGREKLCCCELRLNSSQTQNGNENQGEVEEQTFKEKELDRKPED 1690
QY 457 -----IMRHYLMQKLKNRLKKENKPVIPLPQILGLTASPGVCG----- 495
Db 1691 VPPEILSNERYALQK-ANNRLKILLEVVKTTAAVEETIGHVILGILDRSSKSQSASLI 1749
QY 496 -----ATQKAKAEHILKLCANLDAFTIKTVKENLDQLKNQIQEPCKK-----F 539
Db 1750 WRSEAEASVKSVCVHEEHTRVTDIESPSYSGSDMPRNDINMWSKVTEEGTELSQRLVRSGF 1809
QY 540 AIADATREDPFKEKL-LEIMTRIQTYCQ--MSPMSDFGTQ-----PYEQWAIQMEKKAACK 592
Db 1810 A---GTEIDPENBELMLNISSRLQAAVEKLEIAISSETSSQLEHAKVTQTELMRESFRQKQ 1866
QY 593 GNRKERVCAEHLRK--YNE-----ALQINDTIRMIDAYTHLETFFYNEEKDKKFAVI-- 641
Db 1867 EATESLKQOEELRERLHEESRAREQLAVELSKAEGVIDGYADEKTLFERQIOEKTDIIR 1926
QY 642 -----EDDSDEGGDDE----- 652
Db 1927 LEQELLCASNRLQLEAEQQQIQEERELLRSQKAEAMKAGPVEQQLQETEKLMKXKLE 1986
QY 653 -YCDGDEDEDLKKPLKLDSTD-----RFLMTLFFENNKMMLKRLAENPEYENKLTKLRL 705
Db 1987 VQCAEKVRDLQKQVKALEIDVEEQVSRFI-ELEQEKNTELMDLRQONQALEKQLEKMR 2045
QY 706 NTIMEQYTRTESARGIIFTKTRQSAVALSQWITENEFKFAEVCVKAHHLIGAGHSSEFPKP 765
Db 2046 KFLDEQAIDREHERD-----VFQEQIQKLEQQLKV-----VPRFPQ 2081
QY 766 MTQNEQKEVIS----KFGTKINLLIATTVAEEGL--DIKECNIVIRYGLVTNEIAMVQ 818
Db 2082 ISEHQTRVEQLANHLKEKTKDCSELL---LSKEQLQORDIQERN-----EEIEKLE 2129
QY 819 ARGRARADESTYVLVAHSGS-----GVIE----- 842
Db 2130 FRVR----ELEQALLVSADTFQKVEDRKHFAGAVEAKPELSLEVQLQAERDAIDRKEKEIT 2185
QY 843 --HETVNDFREKNMYKAIHCVQNMPKEBYAHKILELOMQSTMEKKMKTKR-----NIAKH 895
Db 2186 NLEEQLQFREEL-----ENKNEEVQQLHMQLEIQKKESTTRLQELEQENKL 2232
QY 896 YKNPNSLITFLCKNSVLACSGEDIHVI-----EKMHHVN-----MTPEFKELYIV 941
Db 2233 FKDDMEKLGLAIKESD--AMSTQDQHVLPFGKFAQIIQOEKEVEIDQLNEQVTKLQQQLKIT 2290
QY 942 RENKALQKK 950
Db 2291 TDNKVIEEK 2299

RESULT 7

US-08-328-254-6
; Sequence 6, Application US/08328254
; Patent No. 5710022
; GENERAL INFORMATION:
; APPLICANT: Zhu, Xueliang
; APPLICANT: Lee, Wen-Hwa
; TITLE OF INVENTION: A No. 5710022el Nuclear Mitotic Phosphoprotein
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego

; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/328,254
; FILING DATE: 24-OCT-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/141,239
; FILING DATE: 22-OCT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-CJ 1191
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2482 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-328-254-6

Query Match 3.7%; Score 194.5; DB 1; Length 2482;
Best Local Similarity 19.2%; Pred. No. 8.3e-07;
Matches 219; Conservative 186; Mismatches 473; Indels 265; Gaps 46;
QY 21 RVKMYIQVEPVL-----DYTLFLPAEVKEIQIORTVATSG-----NMQAVELLSS 64
Db 1189 KIEACIELEKIVGELKKENSIDLSEKLEYFSCDHQELLQRTVETSEGLNSDLEMHADKSSRE 1248
QY 65 TLEKGVVHL--GWTREFVEALRRGTGSPLAARYMNPETLPLSPSPFENAHDEYLQLLNLQ 122
Db 1249 DIGDNVAKVNDKSWKERFLD-----VENELSRIRSEKASIEH-EALYLEADLE 1294
QY 123 PTLVDKL-LVRDVLDK----CMEEELLTI-EDNRRIAAENNGNE--SGVRELLKRIVQ 173
Db 1295 VVQTEKLCLEKDNENKQVIVCLEEELS VVTSEKQVLTGSDTMSKTTALDQLSKMKKE 1354
QY 174 K---ENWFSAFLNVLROT-----GNNELVQELTGSDCSSESNAEIEENLSQ--VDGPPQVEEQ 223
Db 1355 KTOELESQSECLHCIOVAEAEVKEKTELLQTLSS-SDVSELLKDKTHLOEKQLQSLKDSQ 1413
QY 224 LLSTT-----VQPNLEKEVWGMNNS-----SESSPADSSSVVSES DTSLAEGS----- 266
Db 1414 ALSLTKELENIQIAQLNKEKELLVKESESLOARLSESDYEKLVNSKALEAALVEKGEFAL 1473
QY 267 -VSCLDSESLGHNSNMGSDSGTWGSDSDSDE-----ENVAARASPEPELQRLPYQMEVAQP 318
Db 1474 RLSSTQEEV-HQLRRGIEKLRVRIEADKKQLHIAEKLKRERENDSLKDKVENLERELQ 1532
QY 319 ALEGKNIICLP TSGKTRVAVYIAK-DHLDKKKKASEPGKIVLVNKLVLVEQLFRKEF 377
Db 1533 MSEENQELVILDAENSKAEVETLKTQIEEMARSLKVFELDLVTLRSEKENLTQKIQSKQG 1592
QY 378 QPFLKKWYRVIG-----LSGDTQLKISFPFVVKSCDIIISTAQILENSLLNLN----- 426
Db 1593 Q--LSELDKLLSSFKSLLEEKEQAEIQIKEESK-----TAVEMLQNLKELNEAVALC 1644
QY 427 GEDAGVQLSDFSLIIDECHHTNKEAVYNNIMRHYLMQKLKNRLKKENKPVIPLPQI-- 484
Db 1645 GDQEIMKATEQSLDPPPIEEHQLRNSI-----EKLRLARLEADEKKQLCVLQQLKE 1694
QY 485 -----LGLTASPGVGGATKQAKAEHILKL-----CANL D 514
Db 1695 SEHADLLKGRVENLERELEIARTNQEHAALEAENS KGEVETLKAKIEGTMQSLRGLELD 1754

Db 582 EHGIPTKIIEIRDLSITKESEDDTKBAVQVSYEEMTKBEKDTLLMKLEKEMKDAKA 641
QY 952 ADYQ 955
Db 642 LDPE 645

RESULT 9
US-08-353-700-1
; Sequence 1, Application US/08353700
; Patent No. 5599919
; GENERAL INFORMATION:
; APPLICANT: YEN, TIMOTHY J.
; APPLICANT: RATTNER, JEROME B.
; TITLE OF INVENTION: NUCLEIC ACID ENCODING A
; TITLE OF INVENTION: TRANSIENTLY-EXPRESSED KINETOCHORE PROTEIN,
; TITLE OF INVENTION: AND METHODS OF USE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DANN, DORFMAN, HERRELL AND SKILLMAN
; STREET: 1601 MARKET STREET, SUITE 720
; CITY: PHILADELPHIA
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-2307
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/353,700
; FILING DATE: 09-DEC-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: REED, JANET E.
; REGISTRATION NUMBER: 36,252
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 563-4100
; TELEFAX: (215) 563-4044
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3248 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: HUMAN
; US-08-353-700-1

Query Match 3.6%; Score 191.5; DB 1; Length 3248;
Best Local Similarity 19.2%; Pred. No. 2.3e-06;
Matches 221; Conservative 184; Mismatches 465; Indels 281; Gaps 47;

QY 21 RVKMYIQVEPVL-----DYLTFLPAEVKEIQIQTATSG-----NMQAVELLSS 64
Db 1917 KIEACIELEKIVGELKKENSDDLSEKLEYPFCDHQELLQRVETSEGLNSDLEMHADKSSRE 1976
QY 65 TLEKGVWHL--GWTREFVEALRRRTGSPLAARYMNPETLDLPSPSFENAHDEYLQLNLNQ 122
Db 1977 DIGDNVAKVNDSWKERFLD-----VENELSRIRSEKASIEH-EALYLEADLE 2022
QY 123 PTLVDKL-LVRDVLDK-----CMEEEELTI-EDNRRIAAEENNGNE--SGVRELLKRIVQ 173
Db 2023 VVQTEKLCLEKDNENKQKVVIVCLLEEELSVVTSEKQRLRGELDTMSKTTALDQLSEKME 2082
QY 174 K----ENWFSAFNLVLRQT-----GNNELVQELTGSDCSESABIENTLSQ-VDGPQVEEQ 223
Db 2083 KTOELESQSECLHCIOVAEAEVKEKTELLQTLSDVSELLKDKTHLQEKLOSLEKDSQ 2141

QY 224 LLSTT-----VQPNLEKEVWGMENNS-----SESSFADSSVWSESDTSLAEGS----- 266
Db 2142 ALSLTKELENQIAQLNKEKELLVKESESLQARLSESDYEKLVNSKALEAALVEKGEFAL 2201
QY 267 -VSCLDSESLGHNSNMGSDSGTMCSDSDS-----ENVAARASPEPELQLRPYQMEVAQP 318
Db 2202 RLSSTQEEV-HQLRRGIEKLRVRIEADKKQLHIAEKLKERENDSLKOKVENLERELQ 2260
QY 319 ALEGKNIICLPTGSGKTRVAVVYIAK-DHLDKXKKKASEPGKVIVLVNKVLLVEQLFRKEF 377
Db 2261 MSEENQELVILDAENSKAEVETLKTQIEEMARSLKIFELDLVTLRSEKENLTKQIQEKQG 2320
QY 378 QPFLKKWYRVIG-----LSGDTQLKISFPFVVKSCDIIISTAQILENSLLNLEN----- 426
Db 2321 Q--LSELDKLLSSFKSLLEEKEQABIQIEESK-----TAVEMLQNLKELNEAVALC 2372
QY 427 GEDAGVQLSDFSLIIDECHHTNKEAVYNNIMRHYLMQKLKNRLKKNKPVIPLOI-- 484
Db 2373 GDQEIMKATEQSLDPPIEEHHQLRNSI-----EKLRLARLEADEKKQLCVLQQLKE 2422
QY 485 -----LGLTASPGVGGATKQAKAEHILKL-----CANLD 514
Db 2423 SEHADLLKGRVENLERELEIARTNQEHAALEAENSKGEVETLKAKIEGMTQSLRGLELD 2482
QY 515 AFTIKTVKENLDQLKNQIQEPCKKAFAIADATREDPFKEKLEIMTRIQTICYQMSPMSDFG 574
Db 2483 VVTIRSEKENLTNELQKEQERISELEIINSSFENILQEK----- 2521
QY 575 TQPYEQWAIQMEKAAA-----KKGNRKERVCAEH-----LRKYNEALQIN--- 614
Db 2522 ----EQEKVQMKESSTAMEMLOTQLKELN--ERVAALHNDQEACAKAQNLSSQVECLE 2575
QY 615 -DTIRMIDAYTHLETIFYNEEKDKKFAVIEDSD-----EGSDDEYC---DGDEDEDLKK 665
Db 2576 LEKAQLLQGLDEAKNYYIVLQSSVKGLIQEVEDGKQLEKDEEISRLKNQIQDQEQLVLS 2635
QY 666 PLKLDDETDRFLMTLFFENNMKLKRLAENPEY-----ENEKLTCLRNTIMEQYTRTEES 718
Db 2636 KLSQVEGEH---QLWKEQNLELRNLTVLEQKIQVLQSKNASLQDTLEVQLSSYKNLENE 2692
QY 719 ARGIIFTKTRQSAYALSQWITENEKFAEVGVKAHHLIGAGHSSEFKPMTQNEQKEVISKF 778
Db 2693 ---LELTQMDKMSFV-----BKVNKMTAKETELQREMHEMAQKTAELQEEELSGBKNR 2741
QY 779 RTGKINLL-----IATTVAEEGLDIKECNIVIRYGLVTNEIAMVQARGRA 823
Db 2742 LAGELQLLEEIKSSKDQLKELTLENSELKSLDCMHKQDQVEKEGKVREEIAEYQLRLHE 2801
QY 824 RADESTYVLVAHSGSGVIEHETVNDPREKMMYKAHCVQNMKPE-----EYAHKILE 875
Db 2802 AEKKHQALLDNTNKQYEVEIQT---YREKLTSKE-ECLSSQKLEIDLKSSKEELNNSLK 2857
QY 876 LQMQSIMEKKMKT-----RNIAKHVKNPNPSLITFLCKNCVSLACSGEDIHVIEK--- 925
Db 2858 ATTQ-ILBELKKTMDNLKYVNLKKNENERAQGKMKLLIKSCKQLE---EEKEILQKELS 2913
QY 926 -----MHVNMTPFEFKELYIVRENKA-----LQKKCADYQINGEICK 963
Db 2914 QLQAAQEKQKTGTVMDTKVDELTEIKELKETLEETKEADEYLDKYCS-LLISHEKLEK 2972
QY 964 CGQAWGTMMVH 974
Db 2973 AKEMLETQVAH 2983

RESULT 10
PCT-US95-16216-1
; Sequence 1, Application PC/TUS9516216
; GENERAL INFORMATION:
; APPLICANT: Yen, Timothy J.
; APPLICANT: Rattner, Jerome B.
; TITLE OF INVENTION: Nucleic Acid Encoding a Transiently

```

;
; TITLE OF INVENTION:  Expressed Kinetochore Protein, and Methods of Use
;
; NUMBER OF SEQUENCES:  4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE:  Dann, Dorfman, Herrell and Skillman
; STREET:  1601 Market Street  Suite 720
; CITY:  Philadelphia
; STATE:  PA
; COUNTRY:  USA
; ZIP:  19103-2307
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE:  Floppy disk
; COMPUTER:  IBM PC compatible
; OPERATING SYSTEM:  PC-DOS/MS-DOS
; SOFTWARE:  PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER:  PCT/US95/16216
; FILING DATE:
;
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:  US 08/353,700
; FILING DATE:  09-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME:  Reed, Janet E.
; REGISTRATION NUMBER:  36,252
; TELECOMMUNICATION INFORMATION:
; TELEPHONE:  (215) 563-4100
; TELEFAX:  (215) 563-4044
; INFORMATION FOR SEQ ID NO:  1:
; SEQUENCE CHARACTERISTICS:
; LENGTH:  3248 amino acids
; TYPE:  amino acid
; STRANDEDNESS:  not relevant
; TOPOLOGY:  not relevant
; MOLECULE TYPE:  protein
; HYPOTHETICAL:  NO
; ANTI-SENSE:  NO
; PCT-US95-16216-1

Query Match      3.6%;  Score 191.5;  DB 5;  Length 3248;
Best Local Similarity  19.2%;  Pred. No. 2.3e-06;
Matches 221;  Conservative 184;  Mismatches 465;  Indels 281;  Gaps  47;

QY  21  RVKMYIQVEPVL-----DYLTFPLPAEVKEQIQRTVATSG-----NMQAVELLLS 64
Db  1917  KIEACIELEKIVGELKKENS DLSEKLEYFSCDHQELLQRVETSEGLNSDLEMHADKSSRE 1976

QY  65  TLEKGVWHL--GWTREFVEALRRGTSPLAARYMNPDLTLPSPSPFENAHDEVYQLLNLLQ 122
Db  1977  DIGDNVAKVNDSWKERFLD-----VENELSRIRSEKASIEH-EALYLEADLE 2022

QY  123  PTLVDKL-LVRDVLDK-----CWEEELLTI-EDNRRIAAENNGNE--SGVRELLKRIVQ 173
Db  2023  VVQTEKLCLEKDNEKQKIVICLSEELSVVTSERNQLRGELDTMSKKTALDQLSEKME 2082

QY  174  K-----ENWFSAFLNVLROT-----GNNELVQELTGSDCSSESNAEIEENLSQ-VDGPQVEEQ 223
Db  2083  KTOELESQSECLHCHIQVAEAEVKEKTELLQTLSDVSELLKDKTHLQEKIQSLEKDSQ 2141

QY  224  LLSTT-----VQPNLEKEVWGMENNS-----SESSPADSSVSES DTSILAEGS----- 266
Db  2142  ALSLTKELENQIAQLNKEKELLVKESESLQARLSESDYEKLVNSKALEAALVEKGEFAL 2201

QY  267  -VSCLDSESLGHSNMGSDSGTMSGSDSDE-----ENVAARASPEPELQLRPYQMEVAQP 318
Db  2202  RLSSTQEEV-HQLRRGIEKLRVRIEADKKQLHIAEKLKERERENDSLKDKVENLERELQ 2260

QY  319  ALEGKNIILCLPTGSGKTRVAVYIAK-DHLDKKKKASEPGKVIVLVNKVLLVEQLFRKEF 377
Db  2261  MSEENQELVILDAENSKAEVETLKTQIEEMARSLKIFELDLVTLRSEKENLTQIQIEKQG 2320

QY  378  QPFLKKYRVIG-----LSGDTQLKISFPFVVKSCDIILSTAQILENSLLNLEN----- 426
Db  2321  Q--LSELDKLLSSFKSLLEKEQAEIQIEESK-----TAVEMLQNLKELNEAVALC 2372
```

```

QY  427  GEDAGVQLSDFSLIIIDECHHTNKEAVYNNIMRHVMQKLNRLKKENKVPVLPQI-- 484
      |:  ::  ::  ||  :  ::  ||  :  ::  ||  :  ::  ||  :  ::  ||  :
Db  2373  GDQEIMKATEQSLDPPIEEHQLRNSI-----EKLRLARLEADEKKQLCVLQQLKE 2422
      |:  ::  ::  ||  :  ::  ||  :  ::  ||  :  ::  ||  :  ::  ||  :
QY  485  -----LGLTASPGVGATKQAKAEHHILKL-----CANLD 514
      |:  ::  ::  ||  :  ::  ||  :  ::  ||  :  ::  ||  :  ::  ||  :
Db  2423  SEHHADLLKGRVENLERELEIARTNOEHAALEAENSKGEVETLKAKIEGTMQSLRGLELD 2482
      |:  ::  ::  ||  :  ::  ||  :  ::  ||  :  ::  ||  :  ::  ||  :
QY  515  AFTIKTVKENLDQLKNQIQEPCKKFAIADATREDPFKEKLLLEIMTRIQTQYQMSPMDFG 574
      ||::  ||::  ||:  ::  ||:  ::  ||:  ::  ||:  ::  ||:  ::  ||:
Db  2483  VVTIRSEKENLTNELQEQERISELEIINSFENILQEK----- 2521
      ||::  ||::  ||:  ::  ||:  ::  ||:  ::  ||:  ::  ||:  ::  ||:
QY  575  TPQYEQWAIQMEKKA-----KGNRKERVCAEH-----LRKYNEALQIN-- 614
      ||:  ::  ||:  ::  ||:  ::  ||:  ::  ||:  ::  ||:  ::  ||:
Db  2522  -----EQEKVQMKESSTAMEMLQTOLKELN--ERVAALHNDQEAQKAEQNLSQVECLE 2575
      ||:  ::  ||:  ::  ||:  ::  ||:  ::  ||:  ::  ||:  ::  ||:
QY  615  -DTIRMIDAYTHLETIFYNEEKDKKFAVIEDSD-----EGGDDEYC---DGDEDEDLKK 665
      |:  ::  ::  ||  :  ::  ||  :  ::  ||  :  ::  ||  :  ::  ||  :
Db  2576  LEKAQLLQGLDEAKNNYIVLQSSVKGGLIQEVEDGKQKLEKDEEISRLKNQIQDQEQLVLS 2635
      |:  ::  ::  ||  :  ::  ||  :  ::  ||  :  ::  ||  :  ::  ||  :
QY  666  PLKLDETDRFLMTLFFENNKMKLRLAENPEY-----ENEKLTKLRLNIMEQYTRTEES 718
      |:  ::  ::  ||  :  ::  ||  :  ::  ||  :  ::  ||  :  ::  ||  :
Db  2636  KLSQVEGEH--QLWKEQNLELRLNLTVELEQKIQVLOSKNASLQDTLLEVLOSSYKNLENE 2692
      |:  ::  ::  ||  :  ::  ||  :  ::  ||  :  ::  ||  :  ::  ||  :
QY  719  ARGIIFTKTRQSAYALSQWITENEKFAEVGVKAHHHLIGAGHSSEFKPMTQNEQKEVISKF 778
      |:  ::  ::  ||  :  ::  ||  :  ::  ||  :  ::  ||  :  ::  ||  :
Db  2693  ---LELTKMDKMSFV-----EKVNKMTAKETELQREMHMAQKTAELQOEELSGEKNR 2741
      |:  ::  ::  ||  :  ::  ||  :  ::  ||  :  ::  ||  :  ::  ||  :
QY  779  RTGKINLL-----IATTVAEEGLDIKECNIVIRYGLVTNEIAMVQARGRA 823
      |:  ::  ::  ||  :  ::  ||  :  ::  ||  :  ::  ||  :  ::  ||  :
Db  2742  LAGELQLLLEIISKSKDQLKELTLENSELKSLDCMHKQDQVEKEGVREEIAEYQLRLHE 2801
      |:  ::  ::  ||  :  ::  ||  :  ::  ||  :  ::  ||  :  ::  ||  :
QY  824  RADESTYVLVAHSGSGVIEHETVNDFREKNMYKAIHCVQNMKPE-----EYAHKILE 875
      |:  ::  ::  ||  :  ::  ||  :  ::  ||  :  ::  ||  :  ::  ||  :
Db  2802  AEKKHQALLDITNKOYEVEIQT---YREKLT'SKE-ECLSSQKLEIDLKSSKEELNNSLK 2857
      |:  ::  ::  ||  :  ::  ||  :  ::  ||  :  ::  ||  :  ::  ||  :
QY  876  LMQSIMKKMKT-----RNIAKHYKNPNPSLITFLCKNCSVLACSGEDIHVIK--- 925
      |:  ::  ::  ||  :  ::  ||  :  ::  ||  :  ::  ||  :  ::  ||  :
Db  2858  ATTQ-ILEELKKTMDNLKYVNQLKKENERAQGMKLLIKSKCKQLE--EEKEILQKELS 2913
      |:  ::  ::  ||  :  ::  ||  :  ::  ||  :  ::  ||  :  ::  ||  :
QY  926  -----MHHVNMTPPEFKELYIVRENKA-----LQKKCADYQINGEIIICK 963
      |:  ::  ::  ||  :  ::  ||  :  ::  ||  :  ::  ||  :  ::  ||  :
Db  2914  QLQAAQEKQKTGTVMDTKVDELTTTEIKELKETLEKTKEADEYLDKYCS-LLISHEKLEK 2972
      |:  ::  ::  ||  :  ::  ||  :  ::  ||  :  ::  ||  :  ::  ||  :
QY  964  CGQAWGTMVH 974
      |:  ::  ::  ||  :  ::  ||  :
Db  2973  AKEMLETQVAH 2983
```

```

RESULT 11
US-09-538-092-1154
; Sequence 1154, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT:  Giot, Loic
; APPLICANT:  Mansfield, Traci A.
; TITLE OF INVENTION:  Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE:  15966-542
; CURRENT APPLICATION NUMBER:  US/09/538,092
; CURRENT FILING DATE:  2000-03-29
; PRIOR APPLICATION NUMBER:  60/127,352
; PRIOR FILING DATE:  1999-04-01
; PRIOR APPLICATION NUMBER:  60/178,965
; PRIOR FILING DATE:  2000-02-01
; NUMBER OF SEQ ID NOS:  1387
; SOFTWARE:  CuraPatSeqformatter Version 0.9
; SEQ ID NO 1154
; LENGTH:  3210
; TYPE:  PRT
; ORGANISM:  Homo sapiens
; FEATURE:
; NAME/KEY:  misc_feature
```

; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number P49454
US-09-538-092-1154

Query Match
Best Local Similarity 3.6%; Score 189.5; DB 4; Length 3210;
Matches 218; Conservative 187; Mismatches 473; Indels 265; Gaps 46;

QY 21 RVKMYIQVEPVL-----DYLTFLPAEVKEQIQRTVATSG-----NMQAVELLLS 64
Db 1917 KIEACIELEKIVGELKKNESDLSEKLEYFSCDHQELLQRTVETSEGLNSDLEMHADKSSRE 1976

QY 65 TLEKGVWHL--GWTREFVEALRRTGSPLAARYMNPETDLPSPSFENAHDEYLQLLNLQ 122
Db 1977 DIGDNVAKVNDSWKERFLD-----VENELSRIPSEKASIEH-EALYLEADLE 2022

QY 123 PTLVDKL-LVRDVLDK-----CMEERLLTI-EDNRRIAAENNGNE--SGVRELLKRIVQ 173
Db 2023 VVQTEKLCLEKDNNKQKVVIVCLLEELS VVTSENRQLRGELDTMSKTTALDQLSEKMKKE 2082

QY 174 K----ENWFSAPLNVLRQT-----GNNELVQELTGSDCSNAEINLSQ-VDGPQVVEEQ 223
Db 2083 KTOELES HQSECLHCIOVAEAEVKEKTELLQTLSDVSELLKDKTHLQEKLOSLEKDSQ 2141

QY 224 LLSTT-----VQPNLEKEVGMENNS-----SESSFADSSVSES DTSLAEGS---- 266
Db 2142 ALSLTKELENGIAQLNKEKELLVKESESQARLSESDYEKLVNSKALEAALVEKGEFAL 2201

QY 267 -VSCLDSESLGHSNMGSDSGTMGSDSDE-----ENVAAPASPEPELQLRPYQMEVAQP 318
Db 2202 RLSSTQEEV-HQLRRGIEKLRVRIEADKKQLHIAEKLKERENDSLKDKVENLERELQ 2260

QY 319 ALEGKNIILCLPTGSGKTRVAVYIAK-DHLDKXKKKASEPGKVIVLVNKVLLVEQLFRKEF 377
Db 2261 MSEENQELVLDAENSKAEVETLKTQIEEMARSLKVFELDLVTLRSEKENLTKQIQEKQG 2320

QY 378 QPFLKKWYRIG-----LSGDTQLKISFPPEVVKSCDIIISTAQILENSLLNEN----- 426
Db 2321 Q--LSELDKLLSSFKSLLEEKEQAEIQIKEESK-----TAVEMLQNLKELNEAVALC 2372

QY 427 GEDAGVQLSDPSLIIDECHHTNKEAVYNNIMRHYLMQKLNRLKKNKPVIPQI-- 484
Db 2373 GDQEIMKATEQSLDPPIEEHHQLRNSI-----EKLARLEADEKKQLCVLQQLKE 2422

QY 485 -----LGLTASPGVGGATKQAKAEHHLKL-----CANLD 514
Db 2423 SEHADLLKGRVENLERELEIARTNQEHAALEAENSKGEVETLKAKIEGMTQSLRGLELD 2482

QY 515 AFTIKTVKENLDQLKNQIQEPCKKFAIADATREDPFKEKLLLEIMTRIQTQYQCMSPMSDFG 574
Db 2483 VVTIRSEKEDLTNELQKEQERISELEIINSSFENILQEK-----EQEKVQMKESSTA 2535

QY 575 TQPYEQWAIQMEKKAAGKGNRKERVCA--EHLRKYNEALQIN--DTIRMIDA----YTHL 626
Db 2536 MEMLOTQLKELNERVAALHNDQEAACKAKEQNLSQVECLELEKAQLLQGLDEAKNNYIVL 2595

QY 627 ETFYN-----EEKOKKFAVIEDDSDEGGDDDEYCDGDEDEDLKKPLKIDETD 673
Db 2596 QSSVNGLIQIEVEDGKQKLEKKDEEISRLKNQI-----QDQELVSKLSQVEGE 2643

QY 674 RFLMTLFFENKMKRLAENPEY-----ENEKLTCLRNTIMEQYTRTEESARGIIFTK 726
Db 2644 H---QLWKEQNLERNLTVELEQKIQVLQSKNASLQDTLEVLQSSYKNLENE---LELTK 2697

QY 727 TRQSAYALSQWITENEFKFAEVGVKAHHLIGAGHSSEFFKPMQNEQKEVISKFRGKINLL 786
Db 2698 MDKMSFV-----EKNVMTAKETELQREMHEMAQKTAELQEELSSEKGNRLAGELQLL 2749

QY 787 -----IATTVAEEGLDIKECNIVIRYGLVTNEIAMVQAGRARADESTYV 831
Db 2750 LEEIKSSKDQLKBLTLENSELKSLDCMHKQDQVEKGVKEGBEIEAAYQLRLHEAKKHQAL 2809

QY 832 LVAHSGSGVIEHETVNDFREKMYKAIHCVQNMKPE-----EYAHKILELQMQSIME 883

Db 2810 LLDTNKQYEVEIQT--YREKLTSKE-ECLSSQKLEIDLKSSKEELNNSLKATTQ-ILE 2864

QY 884 KKMKT-----RNIACHYKNNPSLITFLCKNCVSLACSGEDIHVIEK----- 925
Db 2865 ELKKTQMDNLKYVNLKKNENERAQQGMKLLIKSCKQLE---EEKEILOKELSQLQAQBEK 2921

QY 926 -----MHHVNMTPFKELYVRENKA-----LOKKCADIQINGEIIICKGQAWGTM 971
Db 2922 QKTGTVMDTKVDELTTTEIKELKETLEETKEADEYLDKYCS-LLISHEKLEKAKEMLETQ 2980

QY 972 MVH 974
Db 2981 VAH 2983

RESULT 12
US-08-685-576-4
; Sequence 4, Application US/08685576
; Patent No. 5906819
; GENERAL INFORMATION:
; APPLICANT: Kaibuchi, Kozi
; APPLICANT: Iwamatsu, Akihiro
; APPLICANT: Nakano, Takeshi
; APPLICANT: Ito, Masaaki
; APPLICANT: Takahashi, No. 5906819uaki
; TITLE OF INVENTION: RHO TARGET PROTEIN RHO-KINASE
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/685,576
; FILING DATE: 24-JUL-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 7-325129
; FILING DATE: 20-NOV-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 8-17150
; FILING DATE: 05-JAN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 8-131206
; FILING DATE: 26-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Bent, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 16887/843
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1388 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-685-576-4

Query Match 3.5%; Score 188.5; DB 2; Length 1388;
Best Local Similarity 18.2%; Pred. No. 1.1e-06;
Matches 204; Conservative 195; Mismatches 399; Indels 321; Gaps 57;

QY 71 WHLGWTRFVEALRRRTGSPLAARYMNPCL-TDLPSPSFENAHDEYLQLLNLLQPTLVDKL 129
Db 359 WH--W-----DNIRETAAPV-----PELSSDIDSSNFDDIEDD---KGDVETFPKPA 402

QY 130 LVRDVLD-----KCMEEELLTIEDRNRIAAAENNGNESGVRELLKRIVQKENWFSAFNLVL 185
Db 403 FVGNLPPFIFGTYYYRENLLSDSP---SCRENDSIQSRKNEESQEIQKK-----LYTL 452

QY 186 RQTGNNEL-VQELTGSDCSESNAEIEIENLSQVDGPQVEEQLLSTTVQPNLEKEVGMENNS 244
Db 453 EEHLSNEMOAKEELEQCKSVNTRLEKTAK----ELEEEI---TLRKSVESALRQLEREK 505

QY 245 SESSFADSSVSESDTSL-----AEGSVSCLDSESL-----GHNSNMGSDS-GTMGSDSD 292
Db 506 ALLQHKNAEYQORKADHEADKKRNLENDVNSLKQOLEDLKKRNQNSQISTEKNVQLQRQLD 565

QY 293 EENVAARASPEPELQLRPYQMEVAQPA-----LEGKNIIICLPTGSGKTRVAVYIA 343
Db 566 ETNALLRTESDTAARLRKTKQAESSKQIQOLESNNRDLDQKNCL--LETAKLKEKEFINL 623

QY 344 KDHLDKCKKASEPGKVIVLVNKVLLVEQLFRKEFPFLKKWYRVIGLSGDTQLKISFPEV 403
Db 624 QSALESERRDRTHGSEII-----NDLQG-----RICGLEED----- 654

QY 404 VKSCDIIISTAQI----LENSLLNLENGEDAGVQLS-DFSLLII-----DECHH----- 447
Db 655 LKNGKILLAKVELEKRLQERFTDLEK-EKSNMEIDMTYQLKVIQOSLEQEEAEHKATKA 713

QY 448 --TNKEAVYNNIMRHYLMQKLKNRLKKNRPVPLPQILGLTASPGVGGATKQAKAEH 505
Db 714 RLADKNKIYESI-----EEAKSEAMKEMEKKLE-----ERTLKQKVENL 753

QY 506 IL---KLCANLDAFTIKTVKENLDQLKNQIQEPCKKFAIADATREDPFKEKLLEIMTRIQ 562
Db 754 LLEAEKRCSLDC-----DLKOSQOKINELLKQ---KDVLNED-VRNLTLEKIEQETQ 801

QY 563 TYCOMSPMSDFGTQPYEQWAIQMEKKAACKGNRKERVCAEHLRKYNEAL-----QI 613
Db 802 KRC--LTQNDLKMQTQQVNTLKMSEKQLKQENNHLMEMKMNLEKQNAELRKRERQDADQM 859

QY 614 NDTIRMIDAYTHLETFYN-----EEKDK-----KFAVIEDDSD----- 646
Db 860 KELQDQLEABEQYFSTLYKTQVRELKECEEKTLGKELQKQKQELQDERDLSAAQLEITL 919

QY 647 -----EGDDEYCDGDED---EDDLKKPL-----KLDETDRFLMTLFFENNMKL 687
Db 920 TKADSEQLARSIAEEQYSDLEKEKIMKELEIKEMMARHKQELTEKDATIASI-EETNRTL 978

QY 688 KRLAENPEYENEKLTCLRNTIMEQYTRTEE---SARGIIFTKTROSAYALSQWITEN--- 741
Db 979 TSDVANLANEKEELNNKLKDVBQSLSRKDEEISAAI-----KAQFEKQLLTERTLK 1031

QY 742 ----EKPAEYGVKAHHLIGAGHSSEFKPMTQNEQKEVISKFRGTGKINLL----- 786
Db 1032 TQAVNKLAEI-MNRKEPVKRGNDTDVR-RKEKENRKLHMLKSEREKLTOQMIIKYQKELN 1089

QY 787 -IATTVAEGLDIKECNIVIRYGLVTNEIAMVQARGRARA-----DESTYVLVAHSGSV 840
Db 1090 EMQAQIAEE---SQIRIELOMTLDSKSDIEQLRSQALHIGLDSSSI-----GSGP 1139

QY 841 IEHETVNDFREKMM--YKAIHCYQNMKPEEYAHKILELQMQSIMEKKMKTKRNIACHYKN 898
Db 1140 GDAEADDGFPESRLEGWLSLVRNNTKKFGWVKYVIVSSKKILFYDSEQDKE-----QS 1194

QY 899 NPSLITFLCKNSVLACSGEDIHVEIKMHVNMTPPEKEL---YIVRENKALQKKCADY 954
Db 1195 NPYMVLIDIKLFHVRPVTQTDVY-----RADAKEIPRIFQILYANEGESKKEQEF 1244

QY 955 QIN--GE---ICKCGQAWGTMVH-----NCKKQYKKWV 1004
Db 1245 FVEPVEKSNYICHKGHEFIDLYHFTNCEACMKPLWHMFKPPPALECRCHIKCHKOH 1304

QY 975 ----KGLDLPCLKIRNVVVVFKN-----NSTKKQYKKWV 1004

Db 1305 MDKBEIIAPC-KVYYDISTAKNLLLLLANSTEEQ-QKWV 1341

RESULT 13
US-09-976-594-296
; Sequence 296, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 296
; LENGTH: 1388
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6673549 064987CD1
US-09-976-594-296

Query Match 3.5%; Score 188.5; DB 4; Length 1388;
Best Local Similarity 18.2%; Pred. No. 1.1e-06;
Matches 204; Conservative 195; Mismatches 399; Indels 321; Gaps 57;

QY 71 WHLGWTRFVEALRRRTGSPLAARYMNPCL-TDLPSPSFENAHDEYLQLLNLLQPTLVDKL 129
Db 359 WH--W-----DNIRETAAPV-----PELSSDIDSSNFDDIEDD---KGDVETFPKPA 402

QY 130 LVRDVLD-----KCMEEELLTIEDRNRIAAAENNGNESGVRELLKRIVQKENWFSAFNLVL 185
Db 403 FVGNLPPFIFGTYYYRENLLSDSP---SCRENDSIQSRKNEESQEIQKK-----LYTL 452

QY 186 RQTGNNEL-VQELTGSDCSESNAEIEIENLSQVDGPQVEEQLLSTTVQPNLEKEVGMENNS 244
Db 453 EEHLSNEMOAKEELEQCKSVNTRLEKTAK----ELEEEI---TLRKSVESALRQLEREK 505

QY 245 SESSFADSSVSESDTSL-----AEGSVSCLDSESL-----GHNSNMGSDS-GTMGSDSD 292
Db 506 ALLQHKNAEYQORKADHEADKKRNLENDVNSLKQOLEDLKKRNQNSQISTEKNVQLQRQLD 565

QY 293 EENVAARASPEPELQLRPYQMEVAQPA-----LEGKNIIICLPTGSGKTRVAVYIA 343
Db 566 ETNALLRTESDTAARLRKTKQAESSKQIQOLESNNRDLDQKNCL--LETAKLKEKEFINL 623

QY 344 KDHLDKCKKASEPGKVIVLVNKVLLVEQLFRKEFPFLKKWYRVIGLSGDTQLKISFPEV 403
Db 624 QSALESERRDRTHGSEII-----NDLQG-----RICGLEED----- 654

QY 404 VKSCDIIISTAQI----LENSLLNLENGEDAGVQLS-DFSLLII-----DECHH----- 447
Db 655 LKNGKILLAKVELEKRLQERFTDLEK-EKSNMEIDMTYQLKVIQOSLEQEEAEHKATKA 713

QY 448 --TNKEAVYNNIMRHYLMQKLKNRLKKNRPVPLPQILGLTASPGVGGATKQAKAEH 505
Db 714 RLADKNKIYESI-----EEAKSEAMKEMEKKLE-----ERTLKQKVENL 753

QY 506 IL---KLCANLDAFTIKTVKENLDQLKNQIQEPCKKFAIADATREDPFKEKLLEIMTRIQ 562
Db 754 LLEAEKRCSLDC-----DLKOSQOKINELLKQ---KDVLNED-VRNLTLEKIEQETQ 801

QY 563 TYCOMSPMSDFGTQPYEQWAIQMEKKAACKGNRKERVCAEHLRKYNEAL-----QI 613
Db 802 KCC--LTQNDLKMQTQQVNTLKMSEKQLKQENNHLMEMKMNLEKQNAELRKRERQDADQM 859

QY 614 NDTIRMIDAYTHLETFYN-----EEKDK-----KFAVIEDDSD----- 646

Db 860 KELQDLEAEQVSTLYKTQVRELKECEETKLGKELQKQKQELQDERDSLAAQLEITL 919
QY 647 -----EGDDEYCDGDED-----EDDLKKPL-----KLDETDRFLMTLFFENNKM 687
Db 920 TKADSEQLARSTAEQYSDLEKEKIMKELEIKEMMARHKQELTEKDATIASL-EETNRTL 978
QY 688 KRLAENPEYENKLTKLNTIMEQYTRTEE---SARGIIFTKTRQSAYALSQWITEN--- 741
Db 979 TSDVANLANEKEELNNKLKDVQELSRKDEEISAAAI-----KAQFEKQLLTERTLK 1031
QY 742 ----EKFAEVGVKAHHLIGAGHSSEFPKPMTQNEQKEVISKPTGKINLL----- 786
Db 1032 TOAVNKLAEI-MNRKEPVKRGNDTDVR-RKEKENRKLHMLKSEREKLTQOMIKYQKELN 1089
QY 787 -IATTVAEEGLDIKECNIVIRYGLVTNEIAMVQARGRARA-----DESTYVLVAHSGSV 840
Db 1090 EMQAQIAEE---SQIRIELQMTLDSKSDIEQLRSQALHIGLHLDSSSI-----GSGP 1139
QY 841 IEHETVNDFREKMM--YKAIHCVQNMKPEYAHKILELQMSIMEKMKTKRNIKHYKN 898
Db 1140 GDAEADDGFPESRLEGLWSLPVRNNTKKFGWVKYVIVSSKKILFYDSEQDKE-----QS 1194
QY 899 NPSLITFLCKNSVLACSGEDIHVIEKMHVNMTPFEKEL-----YIVRENKALQKKADY 954
Db 1195 NPYMVLIDIDKLFHVRPVTQTDVY-----RADAKEIPRIFQILYANEGESKKBQEF 1244
QY 955 QIN--GE---IICKCGQAWGTMMVH----- 974
Db 1245 PVEPVGEKSNYICHKGHEFIPTLYHFPNCEACMKPLWHMFKPPPALECCRCHIKCHKDH 1304
QY 975 ----KGLDPLCKIRNFVVFKN-----NSTKKQYKKWV 1004
Db 1305 MDKKEIIAPC-KVYVDISTAKNLLLLANSTEEQ-QKWV 1341

RESULT 14
US-09-919-172-29
; Sequence 29, Application US/09919172
; Patent No. 6673545
; GENERAL INFORMATION:
; APPLICANT: Faris, Mary
; APPLICANT: Turner, Christopher M.
; TITLE OF INVENTION: PROSTATE CANCER MARKERS
; FILE REFERENCE: PA-0036 US
; CURRENT APPLICATION NUMBER: US/09/919,172
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/222,469
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PERL Program
; SEQ ID NO 29
; LENGTH: 2125
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673545 3774181CDI
US-09-919-172-29

Query Match 3.5%; Score 187.5; DB 4; Length 2125;
Best Local Similarity 19.6%; Pred. No. 2.6e-06;
Matches 214; Conservative 164; Mismatches 375; Indels 337; Gaps 51;
QY 41 EVKEIQIORTVATSGNMQAVELLLSTLEKGVHGLWTREFVEALRRT----- 86
Db 468 ENQPENSKTATQLNQ--KMLVSEIEMKQSKMDECKYAEQYSATVKDYELQTYRAM 525
QY 87 -----GSPLAARYNPNELTDLPSPSFENAHDEYLLQNLNLLQPTLVDKLLVRDVKCMEE 141
Db 526 VDSQKSPVRRRMQSS-ADLIQEFMDLTRYTALVTLMTQYI---KFAGDSLKRLLEE 581
QY 142 ELLTIED-----RNRIAAEENN---GNESGVRELLKRIVQKENWFS AFLNVLR 186

Db 582 EIKRCKTSEHGAYSDDLQROKATVLENSKLTGKISELERMVAELKKQKSVEEELPKVR 641
QY 187 QTGNNEL-----VQELT-----GSDCSESNAEIENL-SQVDGPQVE-EQLLSTTVQPN 232
Db 642 EAAENELKQQRNVEDISLQKRAESEAQYRRELETIVREKEAAERELERVRLTIEAE 701
QY 233 -----LEKEVWGMENNSESSESPADSSV---VSESDTSLAEGSVSCLDESIGHNSNMGSDS 284
Db 702 AKRAAVEENLLNFRNQLEENTFTRTLEDHLKRKOLSL-----NDLEQQKNKLMEE 752
QY 285 GTMGSDSDEENVAARASPEPELQRPYQMEVAQPALEGKNII----- 326
Db 753 LRRKRDNEBELLKLIKQMEKDL--AFQKQVAEKQLKEKQKIELEARRKITEIQTCREN 809
QY 327 ---ICLPTGSGKTRVAVYIAKDHLDKKKASEPGKIVLVNKKVLLVEQLFRKEFPPLKK 383
Db 810 ALPVCPIQTATSCRAVTGLQOEH--DKQKAE-----LQOQVDELTAANRKAEQDMREL 861
QY 384 WYRVIGLSGDTQLKISFPFVVKSCDIIISTAQILENSLNLNENGEDAGVQLSDPFLIID 443
Db 862 TYELNAL-----QLEKTSSE-----EKARLLKDKL-----D 887
QY 444 ECHHTNKEAVNNIMRHYLMQKLKNNRLKKENKVPILPQILGLTASPGVGGATKQAKAE 503
Db 888 ET-----NNTLR--CLKLELERKQDAEKGYSSQLRELGRQLNQTG-----KAE 929
QY 504 EHILKLCANLDAFTIKTVK-----ENLDQLKNQIQEPCCKFAIADATREDPFK----- 551
Db 930 E-----AMQEASDLKKIKRNYQLELESNLNHEKGLQREVDRITRAHAVAENKIQHLNSQ 983
QY 552 -----EKLLEIMTRIQTQYCOMSPMSDFGTQPYEQWAIQMEK--KAACKGN----- 594
Db 984 IHSFRDEXELE--RLQ-ICQRK--SDHLKEQFEKSHEQLLQNIKAKENNDKIQLNEE 1037
QY 595 -RKERVCAEHL-----RKYN-----EALQI----- 613
Db 1038 LEKSNECAEMLKQKVEELTRONNETKLMQRIQAESENIVLEKQTIQORCEALKIQADGF 1097
QY 614 NDTIRMIDAYTHLETFFYNEEKDKKFAVIEDDSDEGD-----DEYCDGDEDEDLKKPLK 668
Db 1098 KDQLRSTNEHLHKQTKTEQDFQKIKCLEEDLAKSONLVSEBFKQKCD-----QQNII 1149
QY 669 LDETDRFLMTLFFENN--KMLKRLAENP-----EYENEKLTCLR-----TIMEQYT 713
Db 1150 IQNTKKEVRNLNAELNASKEEKRRGEQKVQLQQAQVQELNNRLKKVQDELHLKTIEEQMT 1209
QY 714 R-----TEESARGIIFTKTRQSAYALSQWITENEKFAEVGVKAHHLIGAGHSSEFPKMT 767
Db 1210 HRKMVLFOESG-----KFKQSA-----EEFRK---KMEKLM-----ESKVIT 1244
QY 768 QNEQKEVISKFRGTGKINLLIATTVAEEGLDIKECNIVIRYGLVTNEIAMVQARGRARA 827
Db 1245 END-----ISGIRLDFVSLQQENSRAQENAKLCETNI-----KELERLQOYREQMQQ 1292
QY 828 STYVLVAH-SGSGVIEHETVNDFREKMYKAIHCVQNMKP-----EYAHKILELQMQS 880
Db 1293 GQHMEANYQKQCKLEDELIAQKRE-----VENLKQKMDQIQEHEHQLVLLQCEI 1343
QY 881 IMEKKMKTKRNIKHYKNPNLSLITFLCKNCSVLACSGEDIHVIEKMHVNMTPFEKELYI 940
Db 1344 -----QKKSTAKDCTFKPDF-----EMTVKECQHSSELSSRNTGHLHPTPRSPLLRW 1390
QY 941 VRENKALQKK 950
Db 1391 TQEPQPLEEK 1400

RESULT 15
US-08-973-462-8
; Sequence 8, Application US/08973462B
; Patent No. 6191270
; GENERAL INFORMATION:

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 26, 2005, 20:10:28 ; Search time 162 Seconds
(without alignments)
2285.936 Million cell updates/sec

Title: US-09-515-363C-2
Perfect score: 5311
Sequence: 1 MSGYSTDENFRYLISCFRA.....LPITFPNLDYSECCLFSDDED 1025

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1608061 seqs, 361289386 residues

Total number of hits satisfying chosen parameters: 1608061

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	5311	100.0	1025	14	US-10-055-475-2
2	5311	100.0	1025	14	US-10-228-897-2
3	5311	100.0	1025	16	US-10-755-889-406
4	5307	99.9	1025	14	US-10-055-475-9
5	5299	99.8	1025	14	US-10-055-475-7
6	5285	99.5	1025	15	US-10-275-822A-2
7	2456	46.2	468	14	US-10-094-749-2467
8	1087.5	20.5	925	16	US-10-408-765A-2031
9	1086.5	20.5	925	16	US-10-755-889-234
10	1086.5	20.5	925	17	US-10-370-715B-8
11	1076	20.3	486	14	US-10-055-475-12
12	1040.5	19.6	487	15	US-10-108-260A-4515
13	788.5	14.8	514	14	US-10-055-475-13

14	783	14.7	166	9	US-09-864-761-45372	Sequence 45372, A
15	782	14.7	154	14	US-10-106-698-5088	Sequence 5088, Ap
16	782	14.7	154	15	US-10-264-049-2244	Sequence 2244, Ap
17	617	11.6	416	14	US-10-055-475-14	Sequence 14, Appl
18	563.5	10.6	447	15	US-10-264-049-2303	Sequence 2303, Ap
19	488	9.2	92	9	US-09-864-761-33849	Sequence 33849, A
20	453	8.5	752	14	US-10-369-493-21554	Sequence 21554, A
21	441.5	8.3	650	14	US-10-369-493-1338	Sequence 1338, Ap
22	441.5	8.3	650	14	US-10-369-493-20401	Sequence 20401, A
23	431.5	8.1	1883	16	US-10-437-963-132501	Sequence 132501, A
24	417	7.9	741	14	US-10-369-493-21360	Sequence 21360, A
25	396.5	7.5	485	14	US-10-055-475-17	Sequence 17, Appl
26	395.5	7.4	764	17	US-10-828-924-77	Sequence 77, Appl
27	384	7.2	549	14	US-10-055-475-16	Sequence 16, Appl
28	357	6.7	71	9	US-09-864-761-47821	Sequence 47821, A
29	342	6.4	738	14	US-10-369-493-1227	Sequence 1227, Ap
30	340	6.4	503	14	US-10-055-475-15	Sequence 15, Appl
31	332	6.3	1924	9	US-09-866-557A-2	Sequence 2, Appli
32	332	6.3	1924	11	US-09-858-862-2	Sequence 2, Appli
33	332	6.3	1924	14	US-10-055-797-2	Sequence 2, Appli
34	332	6.3	1924	15	US-10-350-798-2	Sequence 2, Appli
35	310.5	5.8	1557	16	US-10-437-963-137427	Sequence 137427, A
36	298	5.6	55	9	US-09-864-761-48023	Sequence 48023, A
37	296.5	5.6	1583	16	US-10-437-963-159486	Sequence 159486, A
38	295	5.6	1458	16	US-10-437-963-116145	Sequence 116145, A
39	286	5.4	503	13	US-10-027-806-66	Sequence 66, Appl
40	286	5.4	503	13	US-10-034-623-66	Sequence 66, Appl
41	286	5.4	503	14	US-10-027-801-66	Sequence 66, Appl
42	286	5.4	503	14	US-10-029-120-66	Sequence 66, Appl
43	284	5.3	502	13	US-10-027-806-34	Sequence 34, Appl
44	284	5.3	502	13	US-10-034-623-34	Sequence 34, Appl
45	284	5.3	502	14	US-10-027-801-34	Sequence 34, Appl

ALIGNMENTS

RESULT 1
US-10-055-475-2
; Sequence 2, Application US/10055475
; Publication No. US20030022855A1
; GENERAL INFORMATION:
; APPLICANT: Fisher, Paul B.
; APPLICANT: Kang, Dong-Chul
; APPLICANT: Gopalkrishnan, Rahul V.
; TITLE OF INVENTION: USE OF MDA-5 AS AN ANTIVIRAL AND
; TITLE OF INVENTION: ANTIPROLIFERATIVE AGENT
; FILE REFERENCE: A34614-A-PCT-USA-A (070050.1921)
; CURRENT APPLICATION NUMBER: US/10/055,475
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: PCT/US01/06960
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 09/515,363
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1025
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-055-475-2

Query Match 100.0%; Score 5311; DB 14; Length 1025;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1025; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MSGYSTDENFRYLISCFRARKVMYIQVEPVDYLTFLPAEVKEIQRTVATSGNMQAVE	60
Db	1	MSGYSTDENFRYLISCFRARKVMYIQVEPVDYLTFLPAEVKEIQRTVATSGNMQAVE	60
Qy	61	LLSTLEKGVWHLGWTREFVEALRRTGSPLAARYMNPDLTLPSPSFENAHDEYLQLNL	120
Db	61	LLSTLEKGVWHLGWTREFVEALRRTGSPLAARYMNPDLTLPSPSFENAHDEYLQLNL	120

```
QY 121 LQPTLVKLLVRDVLKCMEEELLTIEDNRIAAAENNGNESGVRELLKRIVQENWFS 180
Db 121 LQPTLVKLLVRDVLKCMEEELLTIEDNRIAAAENNGNESGVRELLKRIVQENWFS 180
QY 181 FLNVLRTGNNELVQELTSGDCSESAEINLSQVDPQVVEQLLSTTVQPNLEKEVGM 240
Db 181 FLNVLRTGNNELVQELTSGDCSESAEINLSQVDPQVVEQLLSTTVQPNLEKEVGM 240
QY 241 ENNSSESFADSSVSVSESDTSLAEGSVCLDESGLHNSNMGSDSGTMGSDSDEENVAARA 300
Db 241 ENNSSESFADSSVSVSESDTSLAEGSVCLDESGLHNSNMGSDSGTMGSDSDEENVAARA 300
QY 301 SPEPELQRPYQMEVAQPALEGKNIIICLPTGSGKTRVAVYIAKDHLDKKKASEPGKVI 360
Db 301 SPEPELQRPYQMEVAQPALEGKNIIICLPTGSGKTRVAVYIAKDHLDKKKASEPGKVI 360
QY 361 VLNVKLLVEQLFRKEFPFLKKWYRVIGLSGDTQLKISFPEVVVKSCDIIISTAQILENS 420
Db 361 VLNVKLLVEQLFRKEFPFLKKWYRVIGLSGDTQLKISFPEVVVKSCDIIISTAQILENS 420
QY 421 LLNLENGEDAGVQLSDFSLIIDECHHTNKEAVYNNIMRHYLMOKLKNRLKKNKPVIP 480
Db 421 LLNLENGEDAGVQLSDFSLIIDECHHTNKEAVYNNIMRHYLMOKLKNRLKKNKPVIP 480
QY 481 LPQILGLTASPGVGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQIQEPCCKFA 540
Db 481 LPQILGLTASPGVGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQIQEPCCKFA 540
QY 541 IADATREDPFKEKLEIMTRIQTQYQMSPMDSDFGTQPYEQWAIQMEKKAAGNKRKERV 600
Db 541 IADATREDPFKEKLEIMTRIQTQYQMSPMDSDFGTQPYEQWAIQMEKKAAGNKRKERV 600
QY 601 AEHLRKYNEALQINDTIRMIDAYTHLETIFYNEEKDKKFAVIEDDSDEGGDDEYCDGDEDE 660
Db 601 AEHLRKYNEALQINDTIRMIDAYTHLETIFYNEEKDKKFAVIEDDSDEGGDDEYCDGDEDE 660
QY 661 DDLKKPLKLDDETRFLMTLFFENNMKLKRLAENPEYENEKLTCLRNTIMEQYTRTEESAR 720
Db 661 DDLKKPLKLDDETRFLMTLFFENNMKLKRLAENPEYENEKLTCLRNTIMEQYTRTEESAR 720
QY 721 GIIFTKTQSAYALSQWITENEKFAEVGVKAHHLIGAGHSSEFKPMTQNEQKEVISKFT 780
Db 721 GIIFTKTQSAYALSQWITENEKFAEVGVKAHHLIGAGHSSEFKPMTQNEQKEVISKFT 780
QY 781 GINLLIATTVAEGLDIKECNIVIRYGLVTNEIAMVQARGARADESTYVLVAHSGSV 840
Db 781 GINLLIATTVAEGLDIKECNIVIRYGLVTNEIAMVQARGARADESTYVLVAHSGSV 840
QY 841 IEHETVNDFREKMYKAIHCVQNMKPEEYAHKILELQMSIMEKMKTKRNIKHYKNP 900
Db 841 IEHETVNDFREKMYKAIHCVQNMKPEEYAHKILELQMSIMEKMKTKRNIKHYKNP 900
QY 901 SLITFLCKNCVSLACSGEDIHVIEKMHVNMTPEFKELYIVRENKALQKCADYQINGEI 960
Db 901 SLITFLCKNCVSLACSGEDIHVIEKMHVNMTPEFKELYIVRENKALQKCADYQINGEI 960
QY 961 ICKCGQAWGTMVHKGLDLPCLKIRNFVVVFKNNSTKKQYKKWVELPITFPNLDYSECCL 1020
Db 961 ICKCGQAWGTMVHKGLDLPCLKIRNFVVVFKNNSTKKQYKKWVELPITFPNLDYSECCL 1020
QY 1021 FSDED 1025
Db 1021 FSDED 1025
```

RESULT 2

```
US-10-228-897-2
; Sequence 2, Application US/10228897
; Publication No. US2003092043A1
; GENERAL INFORMATION:
; APPLICANT: Fisher, Paul B.
; APPLICANT: Kang, Dong-Chul
```

```
; APPLICANT: Gopalcrihnan, Rahul V.
; TITLE OF INVENTION: MELANOMA DIFFERENTIATION ASSOCIATED
; FILE REFERENCE: A34614-A-PCT-USA (070050.2121)
; CURRENT APPLICATION NUMBER: US/10/228,897
; CURRENT FILING DATE: 2002-08-26
; PRIOR APPLICATION NUMBER: PCT/US01/06960
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 09/515,363
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1025
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-228-897-2
```

```
Query Match 100.0%; Score 5311; DB 14; Length 1025;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1025; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSGYSTDENFRYLISCPRARVKMYIQVEPVDLYTLFLPAEVKEQIQTATSGNMQAVE 60
Db 1 MSGYSTDENFRYLISCPRARVKMYIQVEPVDLYTLFLPAEVKEQIQTATSGNMQAVE 60
QY 61 LLISTLEKGVHLGWTRFEVREALRRTGSPLAARYMNPDLTDLSPSFENAHDEYQLLNL 120
Db 61 LLISTLEKGVHLGWTRFEVREALRRTGSPLAARYMNPDLTDLSPSFENAHDEYQLLNL 120
QY 121 LQPTLVKLLVRDVLKCMEEELLTIEDNRNIAAAENNGNESGVRELLKRIVQENWFS 180
Db 121 LQPTLVKLLVRDVLKCMEEELLTIEDNRNIAAAENNGNESGVRELLKRIVQENWFS 180
QY 181 FLNVLRTGNNELVQELTSGDCSESAEINLSQVDPQVVEQLLSTTVQPNLEKEVGM 240
Db 181 FLNVLRTGNNELVQELTSGDCSESAEINLSQVDPQVVEQLLSTTVQPNLEKEVGM 240
QY 241 ENNSSESFADSSVSVSESDTSLAEGSVCLDESGLHNSNMGSDSGTMGSDSDEENVAARA 300
Db 241 ENNSSESFADSSVSVSESDTSLAEGSVCLDESGLHNSNMGSDSGTMGSDSDEENVAARA 300
QY 301 SPEPELQRPYQMEVAQPALEGKNIIICLPTGSGKTRVAVYIAKDHLDKKKASEPGKVI 360
Db 301 SPEPELQRPYQMEVAQPALEGKNIIICLPTGSGKTRVAVYIAKDHLDKKKASEPGKVI 360
QY 361 VLNVKLLVEQLFRKEFPFLKKWYRVIGLSGDTQLKISFPEVVVKSCDIIISTAQILENS 420
Db 361 VLNVKLLVEQLFRKEFPFLKKWYRVIGLSGDTQLKISFPEVVVKSCDIIISTAQILENS 420
QY 421 LLNLENGEDAGVQLSDFSLIIDECHHTNKEAVYNNIMRHYLMOKLKNRLKKNKPVIP 480
Db 421 LLNLENGEDAGVQLSDFSLIIDECHHTNKEAVYNNIMRHYLMOKLKNRLKKNKPVIP 480
QY 481 LPQILGLTASPGVGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQIQEPCCKFA 540
Db 481 LPQILGLTASPGVGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQIQEPCCKFA 540
QY 541 IADATREDPFKEKLEIMTRIQTQYQMSPMDSDFGTQPYEQWAIQMEKKAAGNKRKERV 600
Db 541 IADATREDPFKEKLEIMTRIQTQYQMSPMDSDFGTQPYEQWAIQMEKKAAGNKRKERV 600
QY 601 AEHLRKYNEALQINDTIRMIDAYTHLETIFYNEEKDKKFAVIEDDSDEGGDDEYCDGDEDE 660
Db 601 AEHLRKYNEALQINDTIRMIDAYTHLETIFYNEEKDKKFAVIEDDSDEGGDDEYCDGDEDE 660
QY 661 DDLKKPLKLDDETRFLMTLFFENNMKLKRLAENPEYENEKLTCLRNTIMEQYTRTEESAR 720
Db 661 DDLKKPLKLDDETRFLMTLFFENNMKLKRLAENPEYENEKLTCLRNTIMEQYTRTEESAR 720
QY 721 GIIFTKTQSAYALSQWITENEKFAEVGVKAHHLIGAGHSSEFKPMTQNEQKEVISKFT 780
Db 721 GIIFTKTQSAYALSQWITENEKFAEVGVKAHHLIGAGHSSEFKPMTQNEQKEVISKFT 780
```


QY 781 GKINLLIATTVAEEGLDIKECNIVIRYGLVTNEIAMVQARGRARADESTYVLVAHSGSV 840
Db 781 GKINLLIATTVAEEGLDIKECNIVIRYGLVTNEIAMVQARGRARADESTYVLVAHSGSV 840
QY 841 IEHETVNDFREKMMYKAHCVQNMPPEEYAHKILELQMOSIMEKMKTKRNI AKHYKNP 900
Db 841 IEHETVNDFREKMMYKAHCVQNMPPEEYAHKILELQMOSIMEKMKTKRNI AKHYKNP 900
QY 901 SLITFLCKNCSVLACSGEDIHVIEKMHVNMTPEFKELYIVRENKALQKKCADIYQINGEI 960
Db 901 SLITFLCKNCSVLACSGEDIHVIEKMHVNMTPEFKELYIVRENKALQKKCADIYQINGEI 960
QY 961 ICKCGQAWGTMVHKGLDLPCLKIRNFVVVFKNNSTKKQYKKWVLPITFPNLDYSECCL 1020
Db 961 ICKCGQAWGTMVHKGLDLPCLKIRNFVVVFKNNSTKKQYKKWVLPITFPNLDYSECCL 1020
QY 1021 FSDED 1025
Db 1021 FSDED 1025

RESULT 3
US-10-755-889-406
; Sequence 406, Application US/10755889
; Publication No. US20040171823A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-kB
; FILE REFERENCE: D0284 NP
; CURRENT APPLICATION NUMBER: US/10/755,889
; CURRENT FILING DATE: 2004-01-13
; PRIOR APPLICATION NUMBER: U.S. 60/440,068
; PRIOR FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: U.S. 60/469,757
; PRIOR FILING DATE: 2003-05-12
; NUMBER OF SEQ ID NOS: 823
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 406
; LENGTH: 1025
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-755-889-406

Query Match 100.0%; Score 5311; DB 16; Length 1025;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1025; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSGYGSTDENFRYLISCFRARKVMIQVEPVDYLTFLPAEVKEIQIORTVATSGNMQAVE 60
Db 1 MSGYGSTDENFRYLISCFRARKVMIQVEPVDYLTFLPAEVKEIQIORTVATSGNMQAVE 60
QY 61 LLLSTLEKGVWHLGWTFEFVEALRRTGSPLAARYMNPDLTPSPSFENAHDEYLQLNL 120
Db 61 LLLSTLEKGVWHLGWTFEFVEALRRTGSPLAARYMNPDLTPSPSFENAHDEYLQLNL 120
QY 121 LQPTLVDKLLVRDVLDKMEEELLTIEDNRPIAAENNGNESGVRELLKRIVOKENWESA 180
Db 121 LQPTLVDKLLVRDVLDKMEEELLTIEDNRPIAAENNGNESGVRELLKRIVOKENWESA 180
QY 181 FLNVLRTGNNELVQELTGSDCSESNAEINLSQVDGPQVEEQLLSTTVQPNLEKEVWGM 240
Db 181 FLNVLRTGNNELVQELTGSDCSESNAEINLSQVDGPQVEEQLLSTTVQPNLEKEVWGM 240
QY 241 ENNSSESSFADSSVSVESDTSLAEGSVSCLDSESLGHNSNMGSDSGTMGSDSDEENVAARA 300
Db 241 ENNSSESSFADSSVSVESDTSLAEGSVSCLDSESLGHNSNMGSDSGTMGSDSDEENVAARA 300
QY 301 SPEPELQRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKDHLDKKKASEPGKVI 360
Db 301 SPEPELQRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKDHLDKKKASEPGKVI 360

QY 361 VLVNKVLLVEQLFRKEFQPFLLKKYRVIGLSGDTQLKISFPEVVVKSCDIIISTAQILENS 420
Db 361 VLVNKVLLVEQLFRKEFQPFLLKKYRVIGLSGDTQLKISFPEVVVKSCDIIISTAQILENS 420
QY 421 LLNLENGEDAGVQLSDFSLLIIDDECHHTNKEAVYNNIMRHYLMQKLKNNRLKKNKPKVIP 480
Db 421 LLNLENGEDAGVQLSDFSLLIIDDECHHTNKEAVYNNIMRHYLMQKLKNNRLKKNKPKVIP 480
QY 481 LPQILGLTASPGVGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQIQEPCCKPA 540
Db 481 LPQILGLTASPGVGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQIQEPCCKPA 540
QY 541 IADATREDPFKEKLLIEMTRIQTQCMSPMSDFGTQPYEQWAIQMEKKAAGKGNRKERV 600
Db 541 IADATREDPFKEKLLIEMTRIQTQCMSPMSDFGTQPYEQWAIQMEKKAAGKGNRKERV 600
QY 601 AEHLRKYNEALQINDTIRMIDAYTHLETFFYNEEKDKKFAVIEDDSDEGGDDEYCDGDEDE 660
Db 601 AEHLRKYNEALQINDTIRMIDAYTHLETFFYNEEKDKKFAVIEDDSDEGGDDEYCDGDEDE 660
QY 661 DDLKKPLKLDDETDRFLMTLFFENNKMQLKRLAENPEYENEKLTCLRNTIMEQYTRTESAR 720
Db 661 DDLKKPLKLDDETDRFLMTLFFENNKMQLKRLAENPEYENEKLTCLRNTIMEQYTRTESAR 720
QY 721 GIIFTKTROSAYALSQWITENEKFAEVGVKAHHLIGAGHSSEFKPMTQNEQKEVISKPRT 780
Db 721 GIIFTKTROSAYALSQWITENEKFAEVGVKAHHLIGAGHSSEFKPMTQNEQKEVISKPRT 780
QY 781 GKINLLIATTVAEEGLDIKECNIVIRYGLVTNEIAMVQARGRARADESTYVLVAHSGSV 840
Db 781 GKINLLIATTVAEEGLDIKECNIVIRYGLVTNEIAMVQARGRARADESTYVLVAHSGSV 840
QY 841 IEHETVNDFREKMMYKAHCVQNMPPEEYAHKILELQMOSIMEKMKTKRNI AKHYKNP 900
Db 841 IEHETVNDFREKMMYKAHCVQNMPPEEYAHKILELQMOSIMEKMKTKRNI AKHYKNP 900
QY 901 SLITFLCKNCSVLACSGEDIHVIEKMHVNMTPEFKELYIVRENKALQKKCADIYQINGEI 960
Db 901 SLITFLCKNCSVLACSGEDIHVIEKMHVNMTPEFKELYIVRENKALQKKCADIYQINGEI 960
QY 961 ICKCGQAWGTMVHKGLDLPCLKIRNFVVVFKNNSTKKQYKKWVLPITFPNLDYSECCL 1020
Db 961 ICKCGQAWGTMVHKGLDLPCLKIRNFVVVFKNNSTKKQYKKWVLPITFPNLDYSECCL 1020
QY 1021 FSDED 1025
Db 1021 FSDED 1025

RESULT 4
US-10-055-475-9
; Sequence 9, Application US/10055475
; Publication No. US20030022855A1
; GENERAL INFORMATION:
; APPLICANT: Fisher, Paul B.
; APPLICANT: Kang, Dong-Chul
; APPLICANT: Gopalkrishnan, Rahul V.
; TITLE OF INVENTION: USE OF MDA-5 AS AN ANTIVIRAL AND
; TITLE OF INVENTION: ANTIPROLIFERATIVE AGENT
; FILE REFERENCE: A34614-A-PCT-USA-A (070050.1921)
; CURRENT APPLICATION NUMBER: US/10/055,475
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: PCT/US01/06960
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 09/515,363
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 1025
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-055-475-9

Query Match		99.9%;	Score 5307;	DB 14;	Length 1025;				
Best Local Similarity		99.9%;	Pred. No. 0;						
Matches 1024;		Conservative	1;	Mismatches	0;	Indels	0;	Gaps	0;
QY	1	MSGYSTDENFRYLISCFRARKMYIQVEPVDYLTFLPAEVEKIQIORTVATSGNMQAVE	60						
Db	1	MSGYSTDENFRYLISCFRARKMYIQVEPVDYLTFLPAEVEKIQIORTVATSGNMQAVE	60						
QY	61	LLLSTLEKGVWHLGWTREFVEALRRTGSPLAARYMNPDLTLPSPSFENAHDEYLQLLNL	120						
Db	61	LLLSTLEKGVWHLGWTREFVEALRRTGSPLAARYMNPDLTLPSPSFENAHDEYLQLLNL	120						
QY	121	LQPTLVDKLLVRDVLDKMEEELLTIEDNRNIAAENNGNESGVRELLKRIVQKENWFS	180						
Db	121	LQPTLVDKLLVRDVLDKMEEELLTIEDNRNIAAENNGNESGVRELLKRIVQKENWFS	180						
QY	181	FLNVLRQTGNNELVQELTGSDCSESNAEIEENLSQVDGPQVEEQLLSTTVQPNLEKEVWGM	240						
Db	181	FLNVLRQTGNNELVQELTGSDCSESNAEIEENLSQVDGPQVEEQLLSTTVQPNLEKEVWGM	240						
QY	241	ENNSSESSFADSSVSVSESDTSLAEGSVCLDESGLHNSNMGSDSGTMGSDSDEENVAARA	300						
Db	241	ENNSSESSFADSSVSVSESDTSLAEGSVCLDESGLHNSNMGSDSGTMGSDSDEENVAARA	300						
QY	301	SPEPELQLRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKHDLKKKASEPGKVI	360						
Db	301	SPEPELQLRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKHDLKKKASEPGKVI	360						
QY	361	VLNVKLVLLVEQLFRKEFQPLKKWYRVIGLSGDTQLKISFPEVVKSCDIIISTAQILENS	420						
Db	361	VLNVKLVLLVEQLFRKEFQPLKKWYRVIGLSGDTQLKISFPEVVKSCDIIISTAQILENS	420						
QY	421	LLNLENGEDAGVQLSDFSIIIDECHHTNKEAVYNNIMRHYLMQKLKNRLKKNKPVIP	480						
Db	421	LLNLENGEDAGVQLSDFSIIIDECHHTNKEAVYNNIMRHYLMQKLKNRLKKNKPVIP	480						
QY	481	LPQILGLTASPGVGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQIQEPCCKFA	540						
Db	481	LPQILGLTASPGVGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQIQEPCCKFA	540						
QY	541	ICKCGQAWGTMMVHKGLDLPCLKIRNFVVVFKNNSTKKQYKKWVLPITFPNLDYSECCL	1020						
Db	961	ICKCGQAWGTMMVHKGLDLPCLKIRNFVVVFKNNSTKKQYKKWVLPITFPNLDYSECCL	1020						

QY	1021	FSDED 1025							
Db	1021	FSDED 1025							
RESULT 5									
US-10-055-475-7									
; Sequence 7, Application US/10055475									
; Publication No. US20030022855A1									
; GENERAL INFORMATION:									
; APPLICANT: Fisher, Paul B.									
; APPLICANT: Kang, Dong-Chul									
; APPLICANT: Gopalkrishnan, Rahul V.									
; TITLE OF INVENTION: USE OF MDA-5 AS AN ANTIVIRAL AND									
; TITLE OF INVENTION: ANTIPROLIFERATIVE AGENT									
; FILE REFERENCE: A34614-A-PCT-USA-A (070050.1921)									
; CURRENT APPLICATION NUMBER: US/10/055,475									
; CURRENT FILING DATE: 2002-01-22									
; PRIOR APPLICATION NUMBER: PCT/US01/06960									
; PRIOR FILING DATE: 2001-02-28									
; PRIOR APPLICATION NUMBER: 09/515,363									
; PRIOR FILING DATE: 2000-02-29									
; NUMBER OF SEQ ID NOS: 17									
; SOFTWARE: FastSeq for Windows Version 4.0									
; SEQ ID NO 7									
; LENGTH: 1025									
; TYPE: PRT									
; ORGANISM: homo sapiens									
US-10-055-475-7									
Query Match 99.8%; Score 5299; DB 14; Length 1025;									
Best Local Similarity 99.8%; Pred. No. 0;									
Matches 1023; Conservative 1; Mismatches 1; Indels 0; Gaps 0;									
QY	1	MSGYSTDENFRYLISCFRARKMYIQVEPVDYLTFLPAEVEKIQIORTVATSGNMQAVE	60						
Db	1	MSGYSTDENFRYLISCFRARKMYIQVEPVDYLTFLPAEVEKIQIORTVATSGNMQAVE	60						
QY	61	LLLSTLEKGVWHLGWTREFVEALRRTGSPLAARYMNPDLTLPSPSFENAHDEYLQLLNL	120						
Db	61	LLLSTLEKGVWHLGWTREFVEALRRTGSPLAARYMNPDLTLPSPSFENAHDEYLQLLNL	120						
QY	121	LQPTLVDKLLVRDVLDKMEEELLTIEDNRNIAAENNGNESGVRELLKRIVQKENWFS	180						
Db	121	LQPTLVDKLLVRDVLDKMEEELLTIEDNRNIAAENNGNESGVRELLKRIVQKENWFS	180						
QY	181	FLNVLRQTGNNELVQELTGSDCSESNAEIEENLSQVDGPQVEEQLLSTTVQPNLEKEVWGM	240						
Db	181	FLNVLRQTGNNELVQELTGSDCSESNAEIEENLSQVDGPQVEEQLLSTTVQPNLEKEVWGM	240						
QY	241	ENNSSESSFADSSVSVSESDTSLAEGSVCLDESGLHNSNMGSDSGTMGSDSDEENVAARA	300						
Db	241	ENNSSESSFADSSVSVSESDTSLAEGSVCLDESGLHNSNMGSDSGTMGSDSDEENVAARA	300						
QY	301	SPEPELQLRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKHDLKKKASEPGKVI	360						
Db	301	SPEPELQLRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKHDLKKKASEPGKVI	360						
QY	361	VLNVKLVLLVEQLFRKEFQPLKKWYRVIGLSGDTQLKISFPEVVKSCDIIISTAQILENS	420						
Db	361	VLNVKLVLLVEQLFRKEFQPLKKWYRVIGLSGDTQLKISFPEVVKSCDIIISTAQILENS	420						
QY	421	LLNLENGEDAGVQLSDFSIIIDECHHTNKEAVYNNIMRHYLMQKLKNRLKKNKPVIP	480						
Db	421	LLNLENGEDAGVQLSDFSIIIDECHHTNKEAVYNNIMRHYLMQKLKNRLKKNKPVIP	480						
QY	481	LPQILGLTASPGVGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQIQEPCCKFA	540						
Db	481	LPQILGLTASPGVGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQIQEPCCKFA	540						
QY	541	IADATREDPFKEKLLLEINTRIQTQYCOMSPMSDFGTQPYEQWAIQMEKKAACKGNRKERV	600						
Db	541	IADATREDPFKEKLLLEINTRIQTQYCOMSPMSDFGTQPYEQWAIQMEKKAACKGNRKERV	600						

Qy 601 AEHLRKYNEALQINDTIRMIDAYTHLETIFYNEEKDKKFAVIEDSDSDEGGDEYCDGDEDE 660
Db 601 AEHLRKYNEALQINDTIRMIDAYTHLETIFYNEEKDKKFAVIEDSDSDEGGDEYCDGDEDE 660
Qy 661 DDLKKPLKLDLTDRLMTLFFENNKMRLAENPEYENEKLTCLRNTIMEQYTRTEESAR 720
Db 661 DDLKKPLKLDLTDRLMTLFFENNKMRLAENPEYENEKLTCLRNTIMEQYTRTEESAR 720
Qy 721 GIIFTKTQSAYALSQWITENEKFAEVGVKAHHLIGAGHSSEFPKPMQNEQKEVISKFT 780
Db 721 GIIFTKTQSAYALSQWITENEKFAEVGVKAHHLIGAGHSSEFPKPMQNEQKEVISKFT 780
Qy 781 GKINLLIATTVAEEGLDIKECNIVIRYGLVTNEIAMVQARGRARADESTYVLVAHSGSGV 840
Db 781 GKINLLIATTVAEEGLDIKECNIVIRYGLVTNEIAMVQARGRARADESTYVLVAHSGSGV 840
Qy 841 IEHETVNDFREKQMYKAIHCVQNMKPEEYAHKILELQMQSIMCKMKTKRNIACHYKNNP 900
Db 841 IERETVNDFREKQMYKAIHCVQNMKPEEYAHKILELQMQSIMCKMKTKRNIACHYKNNP 900
Qy 901 SLITFLCKNCVSLACSGEDIHVEIKMHVNMTPEFKELYIVRENKALQKCCADYQINGEI 960
Db 901 SLITFLCKNCVSLACSGEDIHVEIKMHVNMTPEFKELYIVRENKALQKCCADYQINGEI 960
Qy 961 ICKCGQAWGTMVHKGLDLPCLKIRNFVVVFKNSTKKQYKKWVLPITFPNLDYSECCL 1020
Db 961 ICKCGQAWGTMVHKGLDLPCLKIRNFVVVFKNSTKKQYKKWVLPITFPNLDYSECCL 1020
Qy 1021 FSDED 1025
Db 1021 FSDED 1025

RESULT 6
US-10-275-822A-2
; Sequence 2, Application US/10275822A
; Publication No. US20040086500A1
; GENERAL INFORMATION:
; APPLICANT: Bahr, Georges
; APPLICANT: Cocude, Cecile
; APPLICANT: Capron, Andre
; TITLE OF INVENTION: RH16 Polypeptide and its Fragments and Polynucleotides Encoding
; TITLE OF INVENTION: said Polypeptides and Therapeutic Uses
; FILE REFERENCE: 01753-170
; CURRENT APPLICATION NUMBER: US/10/275,822A
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: PCT/FR01/01441
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: FR 00/06,030
; PRIOR FILING DATE: 2000-05-11
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1025
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-275-822A-2

Query Match 99.5%; Score 5285; DB 15; Length 1025;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1020; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
Qy 1 MSGYSTDENFRYLISCFRARKVMIQVEPVDLYTLFLPAEVKEQIQRTVATSGNMQAVE 60
Db 1 MSGYSTDENFRYLISCFRARKVMIQVEPVDLYTLFLPAEVKEQIQRTVATSGNMQAVE 60
Qy 61 LLLSTLEKGVWHLGWTFEVALRRRTGSPLAARYMNPDLTLPSPSFENAHDEYLQLNL 120
Db 61 LLLSTLEKGVWHLGWTFEVALRRRTGSPLAARYMNPDLTLPSPSFENAHDEYLQLNL 120
Qy 121 LQPTLVDKLLVRDVKCMEEELLTIEDNRNIAAENNGNESGVRELLKRIVQKENWFS 180
Db 121 LQPTLVDKLLVRDVKCMEEELLTIEDNRNIAAENNGNESGVRELLKRIVQKENWFS 180

Db 121 LQPTLVDKLLVRDVKCMEEELLTIEDNRNIAAENNGNESGVRELLKRIVQKENWFS 180
Qy 181 FLNVLROTGNNELVQELTGSDCSESNAELENLSQVDPQVVEEQLLSTTVQPNLEKEVWGM 240
Db 181 FLNVLROTGNNELVQELTGSDCSESNAELENLSQVDPQVVEEQLLSTTVQPNLEKEVWGM 240
Qy 241 ENNSSESSFADSSVVSSEDTSLAEGSVSLDESGLHNSNMGSDSGTMGSDSDEENVAARA 300
Db 241 ENNSSESSFADSSVVSSEDTSLAEGSVSLDESGLHNSNMGSDSGTMGSDSDEENVAARA 300
Qy 301 SPEPELQRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKDHLDKKKKASEPGKVI 360
Db 301 SPEPELQRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKDHLDKKKKASEPGKVI 360
Qy 361 VLVNKKVLLVEQLFRKEFQPFLLKKWYRVIGLSGDTQLKISFPFVVKSCDIIISTAQILENS 420
Db 361 VLVNKKVLLVEQLFRKEFQPFLLKKWYRVIGLSGDTQLKISFPFVVKSCDIIISTAQILENS 420
Qy 421 LLNLENGEDAGVQLSDFSIIIDECHHTNKEAVYNNIMRHYLMOKLKNRLKKNKPVIP 480
Db 421 LLNLENGEDAGVQLSDFSIIIDECHHTNKEAVYNNIMRHYLMOKLKNRLKKNKPVIP 480
Qy 481 LPQILGLTASPGVGGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQIQEPCCKFA 540
Db 481 LPQILGLTASPGVGGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQIQEPCCKFA 540
Qy 541 IADATREDPFKEKLLIEMTRIQTQYQMSPMDFGTQPYEQWAIQMEKKAACKGNRKERV 600
Db 541 IADATREDPFKEKLLIEMTRIQTQYQMSPMDFGTQPYEQWAIQMEKKAACKGNRKERV 600
Qy 601 AEHLRKYNEALQINDTIRMIDAYTHLETIFYNEEKDKKFAVIEDSDSDEGGDEYCDGDEDE 660
Db 601 AEHLRKYNEALQINDTIRMIDAYTHLETIFYNEEKDKKFAVIEDSDSDEGGDEYCDGDEDE 660
Qy 661 DDLKKPLKLDLTDRLMTLFFENNKMRLAENPEYENEKLTCLRNTIMEQYTRTEESAR 720
Db 661 DDLKKPLKLDLTDRLMTLFFENNKMRLAENPEYENEKLTCLRNTIMEQYTRTEESAR 720
Qy 721 GIIFTKTQSAYALSQWITENEKFAEVGVKAHHLIGAGHSSEFPKPMQNEQKEVISKFT 780
Db 721 GIIFTKTQSAYALSQWITENEKFAEVGVKAHHLIGAGHSSEFPKPMQNEQKEVISKFT 780
Qy 781 GKINLLIATTVAEEGLDIKECNIVIRYGLVTNEIAMVQARGRARADESTYVLVAHSGSGV 840
Db 781 GKINLLIATTVAEEGLDIKECNIVIRYGLVTNEIAMVQARGRARADESTYVLVAHSGSGV 840
Qy 841 IEHETVNDFREKQMYKAIHCVQNMKPEEYAHKILELQMQSIMCKMKTKRNIACHYKNNP 900
Db 841 IERETVNDFREKQMYKAIHCVQNMKPEEYAHKILELQMQSIMCKMKTKRNIACHYKNNP 900
Qy 901 SLITFLCKNCVSLACSGEDIHVEIKMHVNMTPEFKELYIVRENKALQKCCADYQINGEI 960
Db 901 SLITFLCKNCVSLACSGEDIHVEIKMHVNMTPEFKELYIVRENKALQKCCADYQINGEI 960
Qy 961 ICKCGQAWGTMVHKGLDLPCLKIRNFVVVFKNSTKKQYKKWVLPITFPNLDYSECCL 1020
Db 961 ICKCGQAWGTMVHKGLDLPCLKIRNFVVVFKNSTKKQYKKWVLPITFPNLDYSECCL 1020
Qy 1021 FSDED 1025
Db 1021 FSDED 1025

RESULT 7
US-10-094-749-2467
; Sequence 2467, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI

; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOUYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2467
; LENGTH: 468
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-094-749-2467

Query Match 46.2%; Score 2456; DB 14; Length 468;
Best Local Similarity 99.1%; Pred. No. 4e-156;
Matches 464; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 558 MTRIQTQYCOMSPMSDFGTQPYEQWAIQMEKKAAGNRRKERVCAEHLRKYNEALQINDTI 617
Db 1 MTRIQTQYCOMSPMSDFGTQPYEQWAIQMEKKAAGNRRKERVCAEHLRKYNEALQINDTI 60

Qy 618 RMIDAYTHLETFFYNEEKDKKFAVIEDDSDEGGDEYCDGDEDDLKKPLKLDTRFLM 677
Db 61 RMIDAYTHLETFFYNEEKDKKFAVIEDDSDEGGDEYCDGDEDDLKKPLKLDTRFLM 120

Qy 678 TLFFENNKMRLAENPEYENKLTCLRNTIMEQYTRTEESARGIIFTKTRQSAVALSQW 737
Db 121 TLFFENNKMRLAENPEYENKLTCLRNTIMEQYTRTEESARGIIFTKTRQSAVALSQW 180

Qy 738 ITENEKFAEVGVKAHHLIGAGHSSEFKPMTQNEQKEVISKFRGTGRINLLIATTVAEEGLD 797
Db 181 ITENEKFAEVGVKAHHLIGAGHSSEFKPMTQNEQKEVISKFRGTGRINLLIATTVAEEGLD 240

Qy 798 IKECNIVIRYGLVTNEIAMVQARGRADESTYVLVAHSGSGVIEHETVNDEREKMMVKA 857
Db 241 IKECNIVIRYGLVTNEIAMVQARGRADESTYVLVAHSGSGVIEHETVNDEREKMMVKA 300

Qy 858 IHCQNMPKEEYAHKILELQMSIMEKKMTKPNIAKHYNPNPSLITFLCKNCSVLACSG 917
Db 301 IHCQNMPKEEYAHKILELQMSIMEKKMTKPNIAKHYNPNPSLITFLCKNCSVLACSG 360

Qy 918 EDIHVIEKMHVNMTPEFKELYIVRENKALQKCADYQINGEIIICKGQAWGTMVHKGL 977
Db 361 EDIHVIEKMHVNMTPEFKELYIVRENKALQKCADYQINGEIIICKGQAWGTMVHKGL 420

Qy 978 DLPCLKIRNFVVVFKNNSTKKQYKKWVELPITFPNLDYSECCCLFSD 1025
Db 421 DLPCLKIRNFVVVFKNNSTKKQYKKWVELPITFPNLDYSECCCLFSD 468

RESULT 8
US-10-408-765A-2031
; Sequence 2031, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.

; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2031
; LENGTH: 925
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-2031

Query Match 20.5%; Score 1087.5; DB 16; Length 925;
Best Local Similarity 31.5%; Pred. No. 6.5e-64;
Matches 328; Conservative 166; Mismatches 361; Indels 185; Gaps 37;

Qy 12 RYLISCFRVRKMYIQVEPVDLYLT--FLPAEVEQIQRTVATSGNMQAVELLSTLEKG 69
Db 6 RRSIQAFQDIYIRKTLDPYILSYMAPWFREEV-QYIQAEKNKGPMEATLFLKFLLE- 63

Qy 70 VVHLGWTREFVEALRRTGSPLAARYMNPETLPLSPSFENAH--DEYLQNLNLLQPTLVD 127
Db 64 LQEGWFRGFLDALDHAG--YSGLYEAIESWD-----FKKIEKLEEYRLLKRLQPEFT 116

Qy 128 KLLVRDVLDKMEEELLTIEDNRNRIAAENNGNESGVRELLKRIVQ--KENWFS AFLNVL 185
Db 117 RIIPDTIISD-LSECLINQEEIEILQICSTKGMAGAEKLVCELLRS DKENWPCTL---- 171

Qy 186 RQTGNNELVQELTGDCSESNAEIEENLSQVDGPQVEEQLLSTTVQPNLEKVVWGMEN--- 242
Db 172 -----KLAEKERNKFSELWIVEKGIKDVETEDLEDKMETS DIQ-----IFYQEDPEC 219

Qy 243 -NSESSFADSSVSESDTSLAEGSVSCLDES LGHNSNMGSDSGTGMGSDSDEENVAARAS 301
Db 220 QNLSNSCPPSEV---SDTNL-----Y S 239

Qy 302 PEPELQLRPYQMEVAQPALEGKNIIICLPTGSGKTRVAVYIAKDHLKKKASEPGKVIV 361
Db 240 P---FKPRNYQLELALPAMKGNTIICAPTGCCKTFVSLICEHHL-KKFPQGGKGVVF 295

Qy 362 LVNKVLLVEQ---LFRKEFQPLKKWYRVIGLSGDTQLKISFPPEVVKSCDIIISTAQILE 418
Db 296 FANQIPVYEQKSVFSKYFE--RHGYRVGTGISGATAENVPVEQIVENNDIILTPQILV 352

Qy 419 NSLLNLENGEDAGVQLSDFSLIIDECHHTNKEAVYNNIMRHYLMQKLKNNRLKKNKPV 478
Db 353 N---NLKKGTP--SLSIFLIMFDECHNTSKQHPYNNIMFNLYDQKLGSS----- 399

Qy 479 IPLPQILGLTASPGVGATQAKABEHILKLCANLDAFTIKTVKENLDQLKNQIQEPCKK 538
Db 400 GPLPQVIGLTASVGVGDAKNTDEALDYICKLCASLDASVIATVKHNLLEELEQVYVYKPKF 459

Qy 539 FAIADATREDPFKE-----KLEINTRIQTQYCOMSPMSDFGTQPYEQW 581
Db 460 FRKVESRISDKFKYIIAQLMRDTESLAKRICOLENLSQIQN-----REFGTQYEQW 512

Qy 582 AIQMEKKA-----AKGNRKERV-----AEHLRKYNEALQINDTIRMIDAYTHLETFFYN 631
Db 513 IVTVQKACWVFPMPDKDESRICKALFLYTSHLRKYNDALIISEHARMKDALDYKDFFS 572

Qy 632 EEKDKKFAVIEDDSDEGGDEYCDGDEDDLKKPLKLDTRFLMTLFFENNKMRLA 691
Db 573 NVRAAGF-----DEIEQDL-----TQR-----FEELQELSESVS 601

Qy 692 ENPEYENKLTCLRNTIMEQYTRTEESARGIIFTKTRQSAVALSQWITENEK--FAEVGV 749
Db 602 RDPSNENPKLEDLCFILQEEYHLPETIT-ILFVKTRALVALDKNWIEGNPKLSFLKPGI 660

QY 750 KAHHLIGAGHSSEFKPMTQNEQKEVISKFR-TGKINLLIATTVAEEGLDIKECNIVIRYG 808
Db 661 ----LTGRGKTNTGWTLPQAKCILDAPKASGDHNLIIATSVADGIDIAQCNLVILYE 716
QY 809 LVTNEIAMVQARGARADESTYVLVAHSGSGVIEHETVNDFREKQMYKAHCVQNMKPEE 868
Db 717 YVGNVIMQIOTRGRARGSKCFLT-SNAGVIEKEQINMYKEQMNDLSILRLQTWDEAV 775
QY 869 YAHKILELQMQSIMKMKTKRNIACHY--KNPSLITFLCKNSVLACSGEDIHVIEKM 926
Db 776 FREKILHIQTH---EKFIRDSQEKPKVPDPKENKKL---LCRKCKALACYTADVRVIEEC 829
QY 927 HHVNMTPFEKELYIVRENKALQKKCADYQINGEIIIC---KCGQAWGTMVMVHKGLDLPCLK 983
Db 830 HYTVLGDFAFKECFVSRPHPK-PKQFSSFEKRAKIFCARQNCSDHWGIHVYKTFEIPVIK 888
QY 984 IRNFVVVFKNNSTKKQYKKW 1003
Db 889 IESFVVEDIATGVQTLYSKW 908
RESULT 9
US-10-755-889-234
; Sequence 234, Application US/10755889
; Publication No. US20040171823A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-kB
; TITLE OF INVENTION: PATHWAY
; FILE REFERENCE: D0284 NP
; CURRENT APPLICATION NUMBER: US/10/755,889
; CURRENT FILING DATE: 2004-01-13
; PRIOR APPLICATION NUMBER: U.S. 60/440,068
; PRIOR FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: U.S. 60/469,757
; PRIOR FILING DATE: 2003-05-12
; NUMBER OF SEQ ID NOS: 823
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 234
; LENGTH: 925
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-755-889-234
Query Match 20.5%; Score 1086.5; DB 16; Length 925;
Best Local Similarity 31.2%; Pred. No. 7.6e-64;
Matches 325; Conservative 166; Mismatches 364; Indels 185; Gaps 35;
QY 12 RYLISCFRFRVKNYIQVEPVDLYLT--FLPAEVKEIQIQTVAATSGNMQAVELLSTLEKG 69
Db 6 RRSQAQDYIRKTLDPYIILSYMAPWFREEV-QYIAEKNNKGPMEATFLFKFLLE- 63
QY 70 VWHLGWTRFVEALRRTGSPLAARYMNPETLDPSPSFENAH--DEVQLNLLLOPTLVD 127
Db 64 LQEEGWFRGFLDALDHAG--YSGLYEAIESWD-----FKKIEKLEEYRLLLLKRLQPEFKT 116
QY 128 KLLVRDVLDKCMEHELITIEDRNRIAAENNGNESGVRELLKRIVQ--KENWFSAFNLVL 185
Db 117 RIIPTDIISD-LSECLINQEECEIILQICSTKGMAGAEKLVCELLRSDKENWPCTL----- 171
QY 186 RQTGNELVQELTGSDCSESAEIEIENLSQVDGPQVEEQLLSTTVQPNLEKEVWGMEN--- 242
Db 172 -----KLAEKERNKFSSELWIVEKGIKDVTEDLEDKMETSDIQ-----IFYQEDPEC 219
QY 243 -NSSESSFADSSVSVSESDTSLAEGSVCLDESLGHSNMGSDSGTMSDSDSEENVAAARAS 301
Db 220 QNLSENSCPPEV---SDTNL-----YS 239
QY 302 PEPQLRPYQMEVAQPALEGNIIICLPTGSGKTRVAVYIAKDHLDKKKASEPGKVIV 361
Db 240 P---FKPRNYQLELALPAMKGNIIICAPTGGCTFVSLLLICEHHL-KKFPQOGKGVVF 295

QY 362 LVNKKVLLVEQ---LFRKEFQPFLLKKWYRVIGLSGDTQLKISPFVVKSCDIIISTAQILE 418
Db 296 FANQIPVYEQOKSVFSKYFE---RHGYRVGTGISGATAENVPVEQIVENNDIIILTPQILV 352
QY 419 NSLLNLENGEDAGVQLSDFSIIIDECHHTNKEAVYNNIMRHYLMQKLKNNRLKKNPKV 478
Db 353 N---NLKKGTP--SLSIFTLMFDECHNTSKQHPYNNIMFNYLDQKLGSS----- 399
QY 479 IPLPQILGLTASPGVGATKQAKAEEHILKLCANLDAFTIKTVKENLDQLKNQIQEPCCK 538
Db 400 GPLPQVIGLTASVGVGDAKNTDEALDYICKLCASLDASVATVGHNLEELEQVVKPKQF 459
QY 539 FAIADATREDPFKE-----KLEIMTRIQTQYQMSPMSPMSDFGTQPYEQW 581
Db 460 FRKVESRISDKFYIIAQLMRDTSIAKRICKDLENLSIQN-----REFGTQKYEQW 512
QY 582 AIQMEKKA----AKGGRKERV-----AEHLRKYNEALQINDTIRMIDAYTHLETFYN 631
Db 513 IVTVQKACMVFMQPDKDEESRICKALFLYTSHLRKYNDALISEHARMKDALDYLDKDFPS 572
QY 632 EEKDKKFAVIEDSDDEGGDEYCDGDEDEDLKKPLKLDDETRFLMTLFFENNKMMLKRLA 691
Db 573 NVRAAGFEEIEQD-----LTQRFEEKLQELESVS 601
QY 692 ENPEYENEKLTCLRNTIMEQYTRTESARGIIFTKTRQSAVALSQWITENEK--FAEVGV 749
Db 602 RDPSENENPKLEDLCFILQEEYHLNPETIT-ILFVKTRALVDALKNWIEGNPKLSFLKPGI 660
QY 750 KAHHLIGAGHSSEFKPMTQNEQKEVISKFR-TGKINLLIATTVAEEGLDIKECNIVIRYG 808
Db 661 ----LTGRGKTNTGWTLPQAKCILDAPKASGDHNLIIATSVADGIDIAQCNLVILYE 716
QY 809 LVTNEIAMVQARGARADESTYVLVAHSGSGVIEHETVNDFREKQMYKAHCVQNMKPEE 868
Db 717 YVGNVIMQIOTRGRARGSKCFLT-SNAGVIEKEQINMYKEQMNDLSILRLQTWDEAV 775
QY 869 YAHKILELQMQSIMKMKTKRNIACHY--KNPSLITFLCKNSVLACSGEDIHVIEKM 926
Db 776 FREKILHIQTH---EKFIRDSQEKPKVPDPKENKKL---LCRKCKALACYTADVRVIEEC 829
QY 927 HHVNMTPFEKELYIVRENKALQKKCADYQINGEIIIC---KCGQAWGTMVMVHKGLDLPCLK 983
Db 830 HYTVLGDFAFKECFVSRPHPK-PKQFSSFEKRAKIFCARQNCSDHWGIHVYKTFEIPVIK 888
QY 984 IRNFVVVFKNNSTKKQYKKW 1003
Db 889 IESFVVEDIATGVQTLYSKW 908
RESULT 10
US-10-370-715B-8
; Sequence 8, Application US/10370715B
; Publication No. US20040258678A1
; GENERAL INFORMATION:
; Patin Docket Preview
; APPLICANT: BODARY, SARAH C.
; APPLICANT: CLARK, HILLARY
; APPLICANT: BRISDELL, HUNTE
; APPLICANT: JACKMAN, JANET
; APPLICANT: SCHOENFELD, JILL R.
; APPLICANT: WILLIAMS, P. MICKEY
; APPLICANT: WOOD, WILLIAM I.
; APPLICANT: WU, THOMAS D.
; TITLE OF INVENTION: Compositions and Methods for the Treatment of Immune
; TITLE OF INVENTION: Related Diseases
; FILE REFERENCE: P1948R1-US
; CURRENT APPLICATION NUMBER: US/10/370,715B
; CURRENT FILING DATE: 2003-02-21
; NUMBER OF SEQ ID NOS: 742
; SEQ ID NO 8
; LENGTH: 925
; TYPE: PRT
; ORGANISM: Homo sapien

US-10-370-715B-8

```
Query Match      20.5%; Score 1086.5; DB 17; Length 925;
Best Local Similarity 31.2%; Pred. No. 7.6e-64;
Matches 325; Conservative 166; Mismatches 364; Indels 185; Gaps 35;

QY 12 RYLISCFRVRVYQVPEVLDYLT--FLPAEVKEIQIORTVATSGNMQAVELLSTLEKG 69
Db 6 RRSLOAFQDYIRKTLDPYILSYMAPWFREEV-QYIQAEKNNKGPMEATLFLKFLLE- 63

QY 70 VWHLGWTRFVEALRRTGSPLAARYMNPDLTSPSPFENAH--DEYLQLLNLLQPTLVD 127
Db 64 LQEEGWFRGFLDALDHAG--YSGLYEAIESWD-----FKKIEKLEEYRLLLRQLQPEFKT 116

QY 128 KLLVRDVLDKMEEELLTIEDNRNIAAAENNGNESGVRELLKRIVQ--KENWFS AFLNVL 185
Db 117 RIIPTDIISD-LSECLINQEECEILQICSTKGMWAGAEKLVCELLRSDKENWPKTL--- 171

QY 186 RQTGNNELVQELTSDCSESNAEIEIENLSQVDPQVVEEQLLSTTVQPNLEKEVWGMEN--- 242
Db 172 -----KLALERKNKPFSELWIVEKGIKDVETEDLEDKMETSDIQ-----IFYQEDPEC 219

QY 243 -NSSESSFADSSVVSSEDTSLAEGSVSCLDESIGHNSNMGSDSTWGS DSDSEENVARAS 301
Db 220 QNLSSENSCPPSEV---SDTNL-----Y S 239

QY 302 PEPELQRPYQMEVAQPALEGKNIICLTGSGKTRVAVYIAKDHLDKKKKASEPGKVIV 361
Db 240 P---FKPRNYQLELALPAMKGKNTIICAPTGGCTFVSLLCHEHL-KKFPQGGKGVVF 295

QY 362 LVNKVLLVEQ---LFRKEFQPFLLKKWYRVIGLSGDTQLKISPEVVKSCDIIISTAQILE 418
Db 296 FANQIPVYEQQKSVFSKYFE---RHGYRVGTISGATAENVPVEQIVENNDIIILTPQILV 352

QY 419 NSLLNLENGEDAGVQLSDFSLIIIDECHHTNKEAVYNNIMRHYLMQKLKNNRLKKNKPV 478
Db 353 N---NLKKG TGP--SLSIFTLMI FDECHNTSKQHPYNNMIMFNLDQKLGSS----- 399

QY 479 IPLQIILGLTASPGVGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQIQBPCKK 538
Db 400 GPLPQVIGLTASVGVDAKNTDEALDYICKLCASLDASVIATVKHNL EELEQVVKPQKF 459

QY 539 FAIADATREDPPKE-----KLEIMTRIQTQYQMS PMSDFGTQPYEQW 581
Db 460 FRKVESRISDKFYIIAQLMRDTESLAKRICKDLENLSQIQN-----REFGTQYEQW 512

QY 582 AIQMEKKA-----AKKGNRKERV C-----AEHLRKYNEALQINDTIRMIDAYTHLETFTYN 631
Db 513 IVTVQKACMVQFQPKDEESRICKALFLYTSHLRKYNDALISEHARMKDALDYKDFFS 572

QY 632 EEKDKKFAVIEDSDDEGGDEYCDGDEDEDDLLKPLKLDTRFLMTLFFENNKM LKRLA 691
Db 573 NVRAAGFEEIEQD-----LTQRFEEKLQELSESVS 601

QY 692 ENPEYENELTKLRNTIMEQYTRTERESARGIIFTKTRQSAYALSQWITENK--FAEVGV 749
Db 602 RDPSENENPKLEDLCFILQEEYHLNPEIT-ILFVKTRALVDALKNWIEGNPKLSFLKPGI 660

QY 750 KAHHLIGAGHSSEFKPMTONEQKEVSKFR-TGKINLLIATVAEGLDIKECNIVIRYG 808
Db 661 ----LTGRGKTNQNTGMTLPAQKCILDAFKASGDHNLIIATSVADEGIDIAQC NLVILYE 716

QY 809 LVTNEIAMVQARGARADESTYVLVAHSGSGVIEHETVNDFREKMYKAIHC VQNMKPEE 868
Db 717 YGVNVIKMIQTRGRARGSKCFLLT-SNAGVIEKEQINMYKEKMND S ILRLQTWEAV 775

QY 869 YAHKILELQMSIMEKMKTKRNI AKHY--KNNP SLITFLCKNSVLACSGEDTHVIEKM 926
Db 776 FREKILHIQTH---EKFIRDSQEKPKPVDPDKENKKL---LCRKCKALACYTADRVVIEEC 829

QY 927 HHVNMTPFEKELYIVRENKALQKKCADDYQINGEII C---KCGQAWGTMVHKGLDLPCLK 983
Db 830 HYTVLGDAFKECFVSRPHPK-PKQFSSFEKRAKIFCARQNC SHDWGIHVKYKTFEIPVIK 888
```

```
QY 984 IRNFVVVFKNSTKKQYKKW 1003
Db 889 IESFVVEDIATGVQTLYSKW 908
```

```
RESULT 11
US-10-055-475-12
; Sequence 12, Application US/10055475
; Publication No. US20030022855A1
; GENERAL INFORMATION:
; APPLICANT: Fisher, Paul B.
; APPLICANT: Kang, Dong-Chul
; APPLICANT: Gopalkrishnan, Rahul V.
; TITLE OF INVENTION: USE OF MDA-5 AS AN ANTIVIRAL AND
; TITLE OF INVENTION: ANTIPROLIFERATIVE AGENT
; FILE REFERENCE: A34614-A-PCT-USA-A (070050.1921)
; CURRENT APPLICATION NUMBER: US/10/055,475
; PRIOR FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: PCT/US01/06960
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 09/515,363
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 486
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-055-475-12
```

```
Query Match      20.3%; Score 1076; DB 14; Length 486;
Best Local Similarity 44.3%; Pred. No. 1.5e-63;
Matches 236; Conservative 95; Mismatches 154; Indels 48; Gaps 9;
```

```
QY 306 LQLRPYQMEVAQPALEGKNIICLTGSGKTRVAVYIAKDHLDKKKKASEPGKVIVLVNK 365
Db 1 MELRSYQMEVIMPALEGKNIILWPTGAGKTRAAAYVAKRHLE-----TVDGAKVVVLVNR 56

QY 366 VLLVEQLFRKEFQPFLLKKWYRVIGLSGDTQLKISPEVVKSCDIIISTAQILENSLLNLE 425
Db 57 VHLVTQ-HGEEFRRMLDGRWTVTTLSGDMGPRAGFGLHARCHDILLICTAELLQMALTSPE 115

QY 426 NGEDAGVQLSDFSLIIIDECHHTNKEAVYNNIMRHYLMQKLKNNRLKKNKPVIPLPQIL 485
Db 116 --EEHVELTVFSLIVVDECHHTKDTVYNNVIMSQYLELKLQRAQ-----PLPQVL 164

QY 486 GLTASPGVGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQIQBPCKKFAIADAT 545
Db 165 GLTASPGTGASKLDGAINHVLQCANLDTWCIMSPQNCPPQLQEHSSQPCQYNLCHRR 224

QY 546 REDPFKEKLLIEMTRIQTQYQMS PMS-DFGTQPYEQWAIQMEKKA AKKGNRKERVCAEHL 604
Db 225 SQDFGDLKKLMDQIHDLHLEMPELSRKFGTQMYEQVVKLSEAAALAGLQEQRVYALHL 284

QY 605 RKYNEALQINDTIRMIDAYTHLETFTFYNEEKDKKFAVIEDSDDEGGDEYCDGDEDDLK 664
Db 285 RRYNDALLIHDTVRAVDALAALQDFYHREHVTKTQIL-----C----- 322

QY 665 KPLKLDDETDRFLMTLFFENNKM LKRLAENPEYENELTKLRNTIMEQYTRTERESARGIIF 724
Db 323 -----AERRLLALFDDRKNELAHLATHGP-ENPKLEMLEKILQRFPS-SSNSPRGIIF 373

QY 725 TKTRQSAYALSQWITENKFAEVGVKAHHLIGAGHSSEFKPMTONEQKEVSKFRTGKIN 784
Db 374 TRTRQSAHSLLLWLQOQGLQTVDIRAQLLIGAGNSSQSTHMTQRDQEQEVIKQFDGTLN 433

QY 785 LLTATTVAEGLDIKECNIVIRYGLVTNEIAMVQARGARADESTYVLVAHSG 837
Db 434 LLVATSVAAEGLDIPHNCNVVRYGLLTNEISMVQARGAWADQSVAFVATEG 486
```

RESULT 12

US-10-108-260A-4515
; Sequence 4515, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1el full length cdna
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4515
; LENGTH: 487
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-108-260A-4515

Query Match 19.6%; Score 1040.5; DB 15; Length 487;
Best Local Similarity 41.1%; Pred. No. 3.6e-61;
Matches 234; Conservative 90; Mismatches 161; Indels 85; Gaps 9;

QY 306 LQLRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKDHLDKXKASEPGKVIVLVNK 365
Db 1 MELRSYQWEVIMPALEGKNIICLPTGAGKTRAAAYVAKRHLE----TVDGAKVVVLVNR 56

QY 366 VLLVEQLFRKEFPFLKKWYRVIGLSGDTQLKISFPEVVKSCDIISTAQILENSLLNLE 425
Db 57 VHLVTQ-HGEFFRMLDGRWTVTTLSDGMGPRAGFGHLAR----- 95

QY 426 NGEDAGVQLSDFSLIIDECHHTNKEAVYNNIMRHVLMQKLNRLKKNKPKVIPLPQIL 485
Db 96 -----CHHTHKDTVYNNVIMSQYLELKLQRAQ-----PLPQVL 127

QY 486 GLTASPGVGGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQIQEPCCKFAIADAT 545
Db 128 GLTASPGTGGASKLDGAINHVLQCANLDTWCIMSPQNCPCQLQEHSSQPCCKQYNLCHRR 187

QY 546 REDPFEKELLEIMTRIQTQYCMSPMS-DFGTQPYEQWAIQMEKKAAGKGNRKERVCAEHL 604
Db 188 SQDPFGDLLKLMQDQIHDLHLEMPELSRKFGTQMYEQVVKLSEAAALAGLQEQRVYALHL 247

QY 605 RYNEALQINDTIRMIDAYTHLETFFYNEEKDKKFAVIEDDSDEGGDDEYCDGDEDDLK 664
Db 248 RRYNDALLIHTVRAVDALAAALQDFYHREHVTKTQIL-----C----- 285

QY 665 KPLKLDETDRFLMTLFFENNKMRLKLAENPEYENKLTKLRLNTIMEQYTRTEESARGIIF 724
Db 286 -----AERLLALFDDRKNELAHATHGP-ENPKLEMLEKILQRFQFS-SSNSPRGIIF 336

QY 725 TKTRQSAYALSQWITENEKFAEVGVKAHHLIGAGHSSEFKPMTQNEQKEVISKFRGKIN 784
Db 337 TRTRQSAHSLLLWLQOQGLQTVDIRAQLLIGAGNSSQSSTHMTQDQEQEVIQKFDGTLN 396

QY 785 LLIATTVAEEGLDIKCNIVIRYGLVTNEIAMVQARGRARADESTYVLVAHSGSGVIEHE 844
Db 397 LLVATSVAAEGLDIPHCNVVRYGLLTNEISMVQARGRARADQSVYAFVATEGSRELKRE 456

QY 845 TVNDFREKMYKAIHCYQNMKPEEYAHKIL 874
Db 457 LINEALETLMQAVAAVQKMDQAEYQAKEL 486

RESULT 13
US-10-055-475-13
; Sequence 13, Application US/10055475
; Publication No. US20030022855A1
; GENERAL INFORMATION:
; APPLICANT: Fisher, Paul B.
; APPLICANT: Kang, Dong-Chul
; APPLICANT: Gopalakrishnan, Rahul V.
; TITLE OF INVENTION: USE OF MDA-5 AS AN ANTIVIRAL AND
; TITLE OF INVENTION: ANTIPROLIFERATIVE AGENT
; FILE REFERENCE: A34614-A-PCT-USA-A (070050.1921)

; CURRENT APPLICATION NUMBER: US/10/055,475
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: PCT/US01/06960
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 09/515,363
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 514
; TYPE: PRT
; ORGANISM: sus scrofa
US-10-055-475-13

Query Match 14.8%; Score 788.5; DB 14; Length 514;
Best Local Similarity 37.4%; Pred. No. 3.1e-44;
Matches 210; Conservative 89; Mismatches 178; Indels 85; Gaps 16;

QY 306 LQLRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKDHLDKXKASEPGKVIVLVNK 365
Db 6 LKPRKYQLELALPAQNGKNTIICAPTGCCKTFVSLICEHHL-KKFPGRGRKGVFFAIQ 64

QY 366 VLLVEQ---LFRKEFPFLKKWYRVIGLSGDTQLKISFPEVVKSCDIISTAQILENSLL 422
Db 65 LPVYEQQKSVFSKHFE---RLGYKVAGISGATSDTVCVEQIVENSIIILTPQILVNCIL 121

QY 423 NLENGEDAGV--QLSDFSLIIDECHHTNKEAVYNNIMRHVLMQKLNRLKKNKPKVIP 480
Db 122 N-----GTIPSLSVFTLMIFDECHNTSKQHPYNNVIMFSYLDRLKLGSS-----DS 166

QY 481 LPQILGLTASPGVGGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQIQEPCCKFA 540
Db 167 LPQVIGLTASVGVGDANKKAEATEYICKLCASLDTSVIATVRDNLLEELEEVVYKPKQKFFR 226

QY 541 IADATREDPFEKELLEIMTRIQTQYCO-----MSPMSDFGTQPYEQWAIQ 585
Db 227 KVELRTTDRFKCIISQLMMEIESLAKSIFEELGTITLGLFQIQNSNFGTKQYEQWIVKV 286

QY 586 EKAA-----KGNRKERV-----AEHLRKYNEALQINDTIRMIDAYTHLETFFYNEEKD 635
Db 287 QKECAVFQMPDKDKESRICKALFSYMSHLRIYNDALIINEHARMKDALDYLKOFFRNIRA 346

QY 636 KKFVIEDDSDEGGDDEYCDGDEDDLKPLKLDETDRFLMTLFFENNKMRLKLAENPE 695
Db 347 AGF-----DEIEQDL-----TQR-----FEELQEELESISIDPS 375

QY 696 YENEKLTKLRLNTIMEQYTRTEESARGIIFTKTRQSAYALSQWITENEKFAEVGVKAHHLI 755
Db 376 NENPKLRDLCLFQEEYHNLNPEP-RTILFVKTRALVDALKKIKENPKLS--FLKPSILT 432

QY 756 GAGHSSEFKPMTQNEQKEVISKFRGKIN-LLIATTVAEEGLDIKCNIVIRYGLVTNEI 814
Db 433 GRGKTQNIQNTLPAQKCVLDTFRTDKNKILITTSVADEGIDIAQCNLVILYEVGVNVI 492

QY 815 AMVQARGRARADES-TYVLVAH 835
Db 493 KMIQTRGRGRARGSKCFLLLTAN 514

RESULT 14
US-09-864-761-45372
; Sequence 45372, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23

; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 45372
; LENGTH: 166
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC010876.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.58
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.53
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.53
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.5
; OTHER INFORMATION: EST_HUMAN HIT: AI906351.1, EVALUE 1.00e-42
; OTHER INFORMATION: SWISSPROT HIT: P14294, EVALUE 2.20e+00
US-09-864-761-45372

Query Match 14.7%; Score 783; DB 9; Length 166;
Best Local Similarity 93.3%; Pred. No. 1.4e-44;
Matches 152; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 1 MSGNGYSTDENFRYLISCFRARVKMYIQVEPVLDYLTFLPAEVKEIQRTVATSGNQAVE 60
Db 1 MSGNGYSTDENFRYLISCFRARVKMYIQVEPVLDYLTFLPAEVKEIQRTVATSGNQAVE 60

QY 61 LLLSTLEKGVWHLGWTRFVEALRRTGSPLAARYMNPDLTLPSPSFENAHDEVYLQLNL 120
Db 61 LLLSTLEKGVWHLGWTRFVEALRRTGSPLAARYMNPDLTLPSPSFENAHDEVYLQLNL 120

QY 121 LQPTLVDKLLVRDVLDKCMEELLTIEDRNRIAAENNGNESG 163
Db 121 LQPTLVDKLLVRDVLDKCMEELLTIEDRNRVGVCSGDGASLLG 163

RESULT 15
US-10-106-698-5088
; Sequence 5088, Application US/10106698
; Publication No. US20030109690A1

; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptides
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 5088
; LENGTH: 154
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (11)_
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC FEATURE
; LOCATION: (17)_
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC FEATURE
; LOCATION: (18)_
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC FEATURE
; LOCATION: (23)_
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC FEATURE
; LOCATION: (25)_
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC FEATURE
; LOCATION: (26)_
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC FEATURE
; LOCATION: (28)_
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC FEATURE
; LOCATION: (75)_
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-106-698-5088

Query Match 14.7%; Score 782; DB 14; Length 154;
Best Local Similarity 94.8%; Pred. No. 1.5e-44;
Matches 145; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 873 ILELQMSIMEKKMKTKRNTAKHYKNPNSLITFLCKNCVLAACSGEDIHVIEKMHVNMT 932
Db 2 ILELQMSIXEKKMXXXRNIAHXNXPNSLITFLCKNCVLAACSGEDIHVIEKMHVNMT 61

QY 933 PEFKELYIVRENKALQKKCADYQINGEIIICKCGQAWGTMVHKGLDLPCLKIRNFVVVFK 992
Db 62 PEFKELYIVRENKXLQKKCADYQINGEIIICKCGQAWGTMVHKGLDLPCLKIRNFVVVFK 121

QY 993 NNSTKKQYKKWVELPITFPNLDYSECCLFSD 1025
Db 122 NNSTKKQYKKWVELPITFPNLDYSECCLFSD 154

Search completed: January 26, 2005, 20:26:13
Job time : 171 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 27, 2005, 17:57:32 ; Search time 9394 Seconds
(without alignments)
5159.883 Million cell updates/sec

Title: US-09-515-363C-2
Perfect score: 5311
Sequence: 1 MSGYSTDENFRYLISCFA.....LPITFPNLDYSECCLFSD 1025

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09515363/runat_26012005_152043_6141/app_query.fasta_1.1223
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09515363 @CGN 1 1 6626 @runat_26012005_152043_6141 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

RESULT 1
CQ776470
LOCUS CQ776470 3380 bp DNA
DEFINITION Sequence 156 from Patent EP1394274.
ACCESSION CQ776470
VERSION CQ776470.1 GI:45379860
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Ohtani,N., Sugita,Y., Yamaya,M., Kubo,H., Nagai,H. and Izuhara,K.
TITLE Methods of testing for bronchial asthma or chronic obstructive
pulmonary disease
JOURNAL Patent: EP 1394274-A 156 03-MAR-2004;
Genox Research, Inc. (JP)
FEATURES
source Location/Qualifiers
1..3380
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ALIGNMENTS

Linear PAT 11-MAR-2004
CQ776470 Sequence 156 from Patent EP1394274.
CQ776470
CQ776470.1 GI:45379860
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Ohtani,N., Sugita,Y., Yamaya,M., Kubo,H., Nagai,H. and Izuhara,K.
Methods of testing for bronchial asthma or chronic obstructive
pulmonary disease
Patent: EP 1394274-A 156 03-MAR-2004;
Genox Research, Inc. (JP)
Location/Qualifiers
1..3380
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 3380
Score: '5311.00 Matches: 1025
Percent Similarity: 100.00% Conservative: 0

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5311	100.0	3380	6	CQ776470 Sequence
2	5311	100.0	3380	9	AF095844 Homo sapi
3	5285	99.5	3372	6	AX300832 Sequence
4	5285	99.5	3373	9	AY017378 Homo sapi

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Best Local Similarity: 100.00%		Mismatches: 0	
Query Match: 100.00%		Indels: 0	
DB: 6		Gaps: 0	
US-09-515-363C-2 (1-1025) x CQ776470 (1-3380)			
QY	1	MetSerAsnGlyTyrSerThrAspGluAsnPheArgTyrLeuIleSerCysPheArgAla	20
Db	169	ATGTCGAATGGGTATTCCACAGACGAGAATTTCCGCTATCTCATCTCGTGGTTCAGGGCC	228
QY	21	ArgValLysMetTyrIleGlnValGluProValLeuAspTyrLeuThrPheLeuProAla	40
Db	229	AGGGTGAAAATGTACATCCAGGTGGAGCCTGTGCTGGACTACCTGACCTTTCTGCCTGCA	288
QY	41	GluValLysGluGlnIleGlnArgThrValAlaThrSerGlyAsnMetGlnAlaValGlu	60
Db	289	GAGGTGAAGGAGCAGATTTCAGAGGACAGTCGCCACCTCCGGGAACATGCAGGCAGTTGAA	348
QY	61	LeuLeuLeuSerThrLeuGluLysGlyValTrpHisLeuGlyTrpThrArgGluPheVal	80
Db	349	CTGCTGCTGAGCACCTTGGAGAAAGGAGTCTGGCACCTTGGTTGGACTCGGGAATTCGTG	408
QY	81	GluAlaLeuArgArgThrGlySerProLeuAlaAlaArgTyrMetAsnProGluLeuThr	100
Db	409	GAGGCCCTCCGGAGAACCGGCAGCCCTCTGGCCGCCCTACATGAACCCCTGAGCTCAGC	468
QY	101	AspLeuProSerProSerPheGluAsnAlaHisAspGluTyrLeuGlnLeuLeuAsnLeu	120
Db	469	GACTTGCCCTCTCCATCGTTTGAGAACGCTCATGATGAATATCTCCAACCTGCTGAACCTC	528
QY	121	LeuGlnProThrLeuValAspLysLeuLeuValArgAspValLeuAspLysCysMetGlu	140
Db	529	CTTCAGCCCACTCTGGTGGACAAGCTTCTAGTTAGAGACGCTCTTGGATAAGTGCATGGAG	588
QY	141	GluGluLeuLeuThrIleGluAspArgAsnArgIleAlaAlaGluAsnAsnGlyAsn	160
Db	589	GAGGAACCTGTTGACAAATTGAAGACAGAAACCGGATTGCTGCTGCAGAAAACAATGGAAAT	648
QY	161	GluSerGlyValArgGluLeuLeuLysArgIleValGlnLysGluAsnTrpPheSerAla	180
Db	649	GAATCAGGTGAAGAGAGCTACTAAAAAGGATTGTGCAGAAAGAAACTGGTTCTCTGCA	708
QY	181	PheLeuAsnValLeuArgGlnThrGlyAsnAsnGluLeuValGlnGluLeuThrGlySer	200
Db	709	TTTTCTGAATGTTCTTCGTCAAACAGGAAACAATGAACCTTGTCGAAGAGTTAACAGGCTCT	768
QY	201	AspCysSerGluSerAsnAlaGluIleGluAsnLeuSerGlnValAspGlyProGlnVal	220
Db	769	GATTGCTCAGAAAGCAATGCAGAGATTGAGAAATTTATCACAAGTTGATGGTCTCTCAAGTG	828
QY	221	GluGluGlnLeuLeuSerThrThrValGlnProAsnLeuGluLysGluValTrpGlyMet	240
Db	829	GAAGAGCAACTTCTTTCAACCACAGTTTCAGCCAAATCTGGAGAAGGAGGTCTGGGGCATG	888
QY	241	GluAsnAsnSerSerGluSerSerPheAlaAspSerSerValValSerGluSerAspThr	260
Db	889	GAGAACTAACTCATCAGAATCATCTTTTGAGATTCTTCTGTAGTTTTCAGAAATCAGACACA	948
QY	261	SerLeuAlaGluGlySerValSerCysLeuAspGluSerLeuGlyHisAsnSerAsnMet	280
Db	949	AGTTTGGCAGAGGAAGTGTCTGCTGCTTAGATGAAAGTCTTGACATAACAGCAACATG	1008
QY	281	GlySerAspSerGlyThrMetGlySerAspSerAspGluGluAsnValAlaAlaArgAla	300
Db	1009	GGCAGTGATTTCAGGCACCATGGGAAGTGATTTCAGATGAAGAGAAATGTGGCAGCAAGACA	1068
QY	301	SerProGluProGluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaLeu	320
Db	1069	TCCCCGGAGCCAGAACTCCAGCTCAGGCCCTTACCAAAATGGAAAGTTGCCAGCCAGCCTTG	1128
QY	321	GluGlyLysAsnIleIleCysLeuProThrGlySerGlyLysThrArgValAlaVal	340
Db	1129	GAAGGGAAGAATATCATCATCTGCCTCCCTACAGGGAGTGGAAAAACCCAGAGTGGCTGTT	1188

QY	341	TyrIleAlaLysAspHisLeuAspLysLysLysLysAlaSerGluProGlyLysValIle	360
Db	1189	TACATTGCCAAGGATCACCTTAGACAAGAAAAAAGCATCTGAGCCTGGAAGGTATA	1248
QY	361	ValLeuValAsnLysValLeuLeuValGluGlnLeuPheArgLysGluPheGlnProphe	380
Db	1249	GTTCTTGTCAATAAGGTACTGCTAGTTGAACAGCTCTTCCCAAGGAGTTCCAAACCATTT	1308
QY	381	LeuLysLysTrpTyrArgValIleGlyLeuSerGlyAspThrGlnLeuLysIleSerPhe	400
Db	1309	TTGAAGAAATGGTATCGTGTATTGGATTAAAGTGTATACCACTGAAAAATATCATTT	1368
QY	401	ProGluValValLysSerCysAspIleIleSerThrAlaGlnIleLeuGluAsnSer	420
Db	1369	CCAGAAGTTGTCAAGTCTGTGATATATTATCAGTACAGTCAAAATCCTTGAAAACTCC	1428
QY	421	LeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPheSerLeuIle	440
Db	1429	CTCTTAAACTTGGAAAAATGGAGAAGATGCTGGTGTTCATATTGTCAGACTTTTCCCTCATTT	1488
QY	441	IleIleAspGluCysHisHisThrAsnLysGluAlaValTyrAsnAsnIleMetArgHis	460
Db	1489	ATCATTTGATGAATGTCATCACACCACAAAGAAAGCAGTGTATAATAACATCATGAGGCAT	1548
QY	461	TyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysLysGluAsnLysProValIlePro	480
Db	1549	TATTTGATGCAGAAAGTTGAAAAACAATAGACTCAAGAAAGAAAAACAACCCAGTGATTCCTC	1608
QY	481	LeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyGlyAlaThrLysGlnAla	500
Db	1609	CTTCTCAGATACTGGGACTAACAGCTTCACCTGGTGTGGAGGGGCCACGAAGCAAGCC	1668
QY	501	LysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThrIleLysThr	520
Db	1669	AAAGCTGAAGAACACATTTTAAAACTATGTGCCAATCTTGTATGCTTACTATTAAAACT	1728
QY	521	ValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluProCysLysLysPheAla	540
Db	1729	GTTAAAGAAAAACCTTGATCAACTGAAAAAACCAAAATACAGGAGCCATGCAAGAAGTTTGCC	1788
QY	541	IleAlaAspAlaThrArgGluAspProPheLysGluLysLeuLeuGluIleMetThrArg	560
Db	1789	ATTGCAGATGCAACCAAGAGAAGATCCCATTTAAAGAGAAAACTTCTAGAAATAATGACAAG	1848
QY	561	IleGlnThrTyrCysGlnMetSerProMetSerAspPheGlyThrGlnProTyrGluGln	580
Db	1849	ATTCAAACTTATTGTCAAATGAGTCCAAATGTCCAGATTTTGGAACTCAACCCCTATGAACAA	1908
QY	581	TrpAlaIleGlnMetGluLysLysAlaAlaLysLysGlyAsnArgLysGluArgValCys	600
Db	1909	TGGGCCATTCAAATGGAAAAAAAAGCTGCAAAAAAAGGAAATCGCAAAAGAACGTGTTTGT	1968
QY	601	AlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMetIle	620
Db	1969	GCAGAACATTTTGAGGAAGTACAATGAGGCCCTACAAATTAATGACACAATTCGAATGATA	2028
QY	621	AspAlaTyrThrHisLeuGluThrPheTyrAsnGluLysAspLysLysPheAlaVal	640
Db	2029	GATCGGTATACCTCATCTTGAAACTTTCTATATTAAGAGAAAGATAAGAAAGTTTCAGTC	2088
QY	641	IleGluAspAspSerAspGluGlyAspAspGluTyrCysAspGlyAspGluAspGlu	660
Db	2089	ATAGAAGATGATAGTATGATGAGGGTGGTATGATGATGATTTGTGATGGTATGAGATGAG	2148
QY	661	AspAspLeuLysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeuPhe	680
Db	2149	GATGATTTAAAGAAACCTTTGAAACTGGATGAAACAGATAGATTTTCTCATGACTTTATTT	2208
QY	681	PheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGluLys	700
Db	2209	TTTGAAAAACAATAAAATCTTGAAAAAGGCTGGCTGAAAAACCCAGAAATATGAAAAATGAAAG	2268

QY	701	LeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArgThrGluGluSerAlaArg	720
Db	2269	CTGACCAAAATTAAGAAATACCATAATGGAGCAATATACTAGGACTGAGGAATCAGCACGA	2328
QY	721	GlyIleIlePheThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTrpIleThrGlu	740
Db	2329	GGAATAATCTTTACAAAAAACACGACAGAGTGCATATGCGCTTTCCCAGTGGATTACTGAA	2388
QY	741	AsnGluLysPheAlaGluValGlyValLysAlaHisHisLeuIleGlyAlaGlyHisSer	760
Db	2389	AATGAAAAATTTGCTGAAGTAGGAGTCAAAGCCCACCATCTGATGGAGCTGGACACAGC	2448
QY	761	SerGluPheLysProMetThrGlnAsnGluGlnLysGluValIleSerLysPheArgThr	780
Db	2449	AGTGAGTTCAAACCCCATGACACAGAATGAACAAAAAGAAGTCATTAGTAAATTTCCGACT	2508
QY	781	GlyLysIleAsnLeuLeuIleAlaThrThrValAlaGluGluGlyLeuAspIleLysGlu	800
Db	2509	GGAAAAATCAATCTGCTTATCGCTACACAGTGGCAGAGAAGGTCTGGATATTAAGAA	2568
QY	801	CysAsnIleValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAlaArg	820
Db	2569	TGTAACATTGTTATCCGTTATGGTCTCGTCACCAATGAAATAGCCATGGTCCAGGCCGT	2628
QY	821	GlyArgAlaArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGlyVal	840
Db	2629	GGTCGAGCCAGAGCTGATGAGAGCACCTACGTCCTGGTTGCTCACAGTGGTTCCAGGAGTT	2688
QY	841	IleGluHisGluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHisCys	860
Db	2689	ATCGAACATGAGACAGTTAATGATTTCCGAGAGAAGATGATGTATAAAGCTATACATTGT	2748
QY	861	ValGlnAsnMetLysProGluGluTyrAlaHisLysIleLeuGluLeuGlnMetGlnSer	880
Db	2749	GTTCAAAATATGAACACGAGGAGTATGCTCATAAGATTTTGGAAATTA CAGATGCAAGT	2808
QY	881	IleMetGluLysLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAsnPro	900
Db	2809	ATAATGGAAAAGAAAATGAAAACCAAGAGAAATATTGCCAAGCATTAAGAATAACCCA	2868
QY	901	SerLeuIleThrPheLeuCysLysAsnCysSerValLeuAlaCysSerGlyGluAspIle	920
Db	2869	TCACATAATACTTCTCTTGCAAAAACGCGAGTGCTAGCCTGTTCTGGGGAAGATATC	2928
QY	921	HisValIleGluLysMetHisHisValAsnMetThrProGluPheLysGluLeuTyrIle	940
Db	2929	CATGTAATTGAGAAAATGCATCACGTCATATATGACCCAGAAATTC AAGGAAC TTTTACATT	2988
QY	941	ValArgGluAsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluIle	960
Db	2989	GTAAGAGAAAACAAAGCACTGCAAAAGAAAGTGTGCCGACTATCAAAATAATGGTGAATC	3048
QY	961	IleCysLysCysGlyGlnAlaTrpGlyThrMetMetValHisLysGlyLeuAspLeuPro	980
Db	3049	ATCTGCAAATGTGGCCAGGCTTGGGGAACAATGATGGTGCACAAAGGCTTAGATTGGCT	3108
QY	981	CysLeuLysIleArgAsnPheValValPheLysAsnAsnSerThrLysLysGlnTyr	1000
Db	3109	TGTCTCAAAATAAGGAATTTTGTAGTGGTTTCAAAAAATAATTCAACAAAGAAACAATAC	3168
QY	1001	LysLysTrpValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCysCysLeu	1020
Db	3169	AAAAAGTGGGTAGAAATTACCTATCACATTTCCCAATCTTGACIATTCAGAAATGCTGTTA	3228
QY	1021	PheSerAspGluAsp	1025
Db	3229	TTTAGTGATGAGGAT	3243
RESULT 2			
AF095844			
LOCUS	AF095844	3380 bp	mRNA
DEFINITION	Homo sapiens melanoma differentiation associated protein-5 (MDA5)	linear	PRI 23-JAN-2002
	mRNA, complete cds.		

ACCESSION	AF095844		
VERSION	AF095844.1	GI:11344593	
KEYWORDS	.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1 (bases 1 to 3380)		
AUTHORS	Kang,D.C., Gopalkrishnan,R.V., Wu,Q., Jankowsky,E., Pyle,A.M. and Fisher,P.B.		
TITLE	mda-5: An interferon-inducible putative RNA helicase with double-stranded RNA-dependent ATPase activity and melanoma growth-suppressive properties		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (2), 637-642 (2002)		
MEDLINE	21664412		
PUBMED	11805321		
REFERENCE	2 (bases 1 to 3380)		
AUTHORS	Kang,D.-C. and Fisher,P.B.		
TITLE	Direct Submission		
JOURNAL	Submitted (30-SEP-1998) Urology, Columbia University, P&S, 630 W168th, New York, NY 10032, USA		
FEATURES	Location/Qualifiers		
source	1..3380		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/cell_line="HO-1"		
	/tissue_type="melanoma"		
gene	1..3380		
	/gene="MDA5"		
CDS	169..3246		
	/gene="MDA5"		
	/codon_start=1		
	/product="melanoma differentiation associated protein-5"		
	/protein_id="AAG34368.1"		
	/db_xref="GI:11344594"		
	/translation="MSGYSTDENFRYLISCFRARKMYIQVEPVDYLTFLPAEVKE QIQT VATSGNMQAVELLSTLEKGVHLGWTREFVEALRRTGSPLAARYMNP ELTDL PPSFENAHDEYLQLNLLOPTLVDKLLVRDVLDKMEEELLTIEDRNR IAA AENNGN ESGVRELLKRI VOKENWFS AFLNVLROTGNNELVQELT GSDCSESNAE IENLSQVDGP QVEEQLLSTTVQPNLEKVGWMENNSSSESFADSSVSVSEDTSLAEGSVSCLDES LGH NSNMGSDSGTMGSDSEENVAARASPEPELQRPYQMEVAQPALEGKN I ICLPTGSG KTRVAVYIAKHLDKKKASEPGKVI VLVNKKVLLVEQLFRKEFPFLKKWYRVIGLSG DTQLKISFPEVVKSDI IISTAQI LENSLLNLENGEDAGVQLSDFSL IIDECHHTNK EAVYNIMRHVLMQKLKNRLKKENKPIVLPQILGLTASPGVGATKQAKAE EHLK LCANLDAFTIKTVKENLDQLKNQIQEPCKEFAIADATREDPFKEKLEIMTRIQT YCQ MSPMSDFGTQPYEQWAIQMEKKA AKGNRKERVCAEHLRKYNEALQINDTIRMIDAYT HLETFYNEEKDKKFAVIDSDSDEGDD EYCDGDEDD LKKPLKLD ETRFLMTLFFE NNKMLKRLAENPEYENEKLT KLRNTIMEQYTRTEESARGI IFTKTRQSAYALSQWITE NEKPAEVGVKAHLIGAGHSSEFPKPMTQNEQKVISKFR TKINLLIATTVAAEGLDI KECNIVIRYGLVTNEI AMVQARGARADESTYLVVAHSGSGVIEHETVNDFREKMYK AIHCVQNMKPEEYAHKILELQMQS IMEKKMKTNRNI AKHYKNP SLITFLCKNC SVLA CSGEDIHVIEKMHVNMNTP EFKELYIVRENKALQKKCADYQINGE I ICKGQAGTMM VHKGDLPLCLKIRNFVVVFKNNSTKKQYKKWVLPITFPNL DYSECCLFSD E D"		
ORIGIN			
Alignment Scores:			
Pred. No.:	0	Length:	3380
Score:	5311.00	Matches:	1025
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0
US-09-515-363C-2 (1-1025) x AF095844 (1-3380)			
QY	1	MetSerAsnGlyTyrSerThrAspGluAsnPheArgTyrLeuIleSerCysPheArgAla	20
Db	169	ATGTGGAATGGGTATTCCACAGACGAGAATTTCCGCTATCTCATCTCGTGTTCAGGCCC	228
QY	21	ArgValLysMetTyrIleGlnValGluProValLeuAspTyrLeuThrPheLeuProAla	40
Db	229	AGGGTGAAAAATGTACATCCAGGTGGAGCCTGTGCTGGACTACCTGACCTTTCTGCCTGCA	288

QY 41 GluValLysGluGlnIleGlnArgThrValAlaThrSerGlyAsnMetGlnAlaValGlu 60
 Db 289 GAGGTGAAGGACAGATTTCAGAGGACAGTCGCCACCTCCGGAAACATGCAGGCAGTTGAA 348
 QY 61 LeuLeuLeuSerThrLeuGluLysGlyValTrpHisLeuGlyTrpThrArgGluPheVal 80
 Db 349 CTGCTGCTGAGCACCTTGGAGAAAGGAGTCTGGCACCTTGGTGGACTCGGGAATTCGTG 408
 QY 81 GluAlaLeuArgArgThrGlySerProLeuAlaAlaArgTyrMetAsnProGluLeuThr 100
 Db 409 GAGGCCCTCCGAGAAACCGGACGCCCTCTGGCCGCCCTACATGAACCCCTGAGCTCAGG 468
 QY 101 AspLeuProSerProSerPheGluAsnAlaHisAspGluTyrLeuGlnLeuLeuAsnLeu 120
 Db 469 GACTTGCCCTCTCCATCGTTTGAGAACGCTCATGATGAATATCTCCAACCTGCTGAACCTC 528
 QY 121 LeuGlnProThrLeuValAspLysLeuLeuValArgAspValLeuAspLysCysMetGlu 140
 Db 529 CTTTCAGCCCACTCTGGTGGACAAGCTTCTAGTTAGAGAGCTCTTGGATAAGTGGATGGAG 588
 QY 141 GluGluLeuLeuThrIleGluAspArgAsnArgIleAlaAlaGluAsnAsnGlyAsn 160
 Db 589 GAGGAACCTGTGACAATTGAAGACAGAAACCGGATTGTCTGCTGCAGAAACAATGGAAT 648
 QY 161 GluSerGlyValArgGluLeuLeuLysArgIleValGlnLysGluAsnTrpPheSerAla 180
 Db 649 GAATCAGGTGTAAAGAGAGCTACTAAAAAGGATTGTGCAGAAAGAAAACCTGTTCTCTGCA 708
 QY 181 PheLeuAsnValLeuArgGlnThrGlyAsnAsnGlnLeuValGlnGluLeuThrGlySer 200
 Db 709 TTTCTGAATGTTCTTCGTCAAACAGGAAACAATGAACCTGTCCAAGAGTTAACAGGCTCT 768
 QY 201 AspCysSerGluSerAsnAlaGluIleGluAsnLeuSerGlnValAspGlyProGlnVal 220
 Db 769 GATTGCTCAGAAAGCAATGCAGAGATTGAGAAATTTATCACAAGTTGATGGTCTCAAGTG 828
 QY 221 GluGluGlnLeuLeuSerThrThrValGlnProAsnLeuGluLysGluValTrpGlyMet 240
 Db 829 GAAGAGCAACTTCTTTCACCACAGTTTCAGCCAAATCTGGAGAGGAGGCTGGGGCATG 888
 QY 241 GluAsnAsnSerSerGluSerPheAlaAspSerSerValValSerGluSerAspThr 260
 Db 889 GAGATAACTCATCAGAAATCATCTTTTGCAGATTCTTCTGTAGTTTTCAGAAATCAGACACA 948
 QY 261 SerLeuAlaGluGlySerValSerCysLeuAspGluSerLeuGlyHisAsnSerAsnMet 280
 Db 949 AGTTTGGCAGAAAGGAAGTGTACGTGCTTAGATGAAAGTCTTGGACATAACAGCAACATG 1008
 QY 281 GlySerAspSerGlyThrMetGlySerAspSerAspGluGluAsnValAlaAlaArgAla 300
 Db 1009 GGCAGTGATTCAGGCACCATGGGAAGTGATTTCAGATGAAGAGAAATGTGGCACAAGAGCA 1068
 QY 301 SerProGluProGluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaLeu 320
 Db 1069 TCCCCGAGCCAGAACTCCAGCTCAGGCCTTACCATAATGGAAGTTGCCCGAGCCAGCCTTG 1128
 QY 321 GluGlyLysAsnIleIleCysLeuProThrGlySerGlyLysThrArgValAlaVal 340
 Db 1129 GAAGGGAAGAATATCATCATCTGCCTCCCTACAGGGAGTGGAAAAACCCAGAGTGGCTGTT 1188
 QY 341 TyrIleAlaLysAspHisLeuAspLysLysLysAlaSerGluProGlyLysValIle 360
 Db 1189 TACATTGCCAAGGATCATCTAGACAAAGAAAGAAAGATCTGAGCCTGGAAAAAGTTATA 1248
 QY 361 ValLeuValAsnLysValLeuLeuValGluGlnLeuPheArgLysGluPheGlnProphe 380
 Db 1249 GTTCTGTCAATAAGGTACTGCTAGTTGAACAGCTCTTCGCAAGGAGTTCCAACCATTT 1308
 QY 381 LeuLysLysTrpTyrArgValIleGlyLeuSerGlyAspThrGlnLeuLysIleSerPhe 400
 Db 1309 TTGAAGAAATGGTATCGTGTATTGGATTAAAGTGGTGATACCCCACTGAAAAATATCATTT 1368

QY 401 ProGluValValLysSerCysAspIleIleIleSerThrAlaGlnIleLeuGluAsnSer 420
 Db 1369 CCAGAGATTGTCAAGTCTCTGTGATATTATTATCAGTACAGCTCAAATCCTTGAAAACTCC 1428
 QY 421 LeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPheSerLeuIle 440
 Db 1429 CTCTTAAACTTGGAAAAATGGAGAAAGATGCTGGTGTCAATTGTTCAGACTTTTCCCTCAT 1488
 QY 441 IleIleAspGluCysHisThrAsnLysGluAlaValTyrAsnAsnIleMetArgHis 460
 Db 1489 ATCATTTGATGAATGTCAACACCAACAAAGAGCAGTGTATATAACATCATGAGGCAT 1548
 QY 461 TyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysGluAsnLysProValIlePro 480
 Db 1549 TATTGATGCAGAAGTTGAAAAACAATAGACTCAAGAAAAAGAAAAACAACCAAGTATCCC 1608
 QY 481 LeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyGlyAlaThrLysGlnAla 500
 Db 1609 CTTCTCTCAGATACTGGGACTAACAGCTTCACCTGGTGTGGAGGGGCCACGAAGCAAGCC 1668
 QY 501 LysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThrIleLysThr 520
 Db 1669 AAAGCTGAAGAACAACATTTTAAACTATGTGCCAATCTTGATGCATTTACTATTAAACT 1728
 QY 521 ValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluProCysLysLysPheAla 540
 Db 1729 GTTAAAGAAAAACCTTGATCAACTGAAAAACCAATAACAGGAGCCATGCAAGAAGTTTGCC 1788
 QY 541 IleAlaAspAlaThrArgGluAspPropheLysGluLysLeuLeuGluIleMetThrArg 560
 Db 1789 ATTGCAGATGCAACCCAGAGAAGATCCATTTAAAGAGAAAACTTCTAGAAATAAATGACAA 1848
 QY 561 IleGlnThrTyrCysGlnMetSerProMetSerAspPheGlyThrGlnProTyrGluGln 580
 Db 1849 ATTCAAACTTATTGTCAAATGAGTCCAATGTCCAGATTTTGGAACTCAACCTTATGAACA 1908
 QY 581 TrpAlaIleGlnMetGluLysLysAlaAlaLysLysGlyAsnArgLysGluArgValCys 600
 Db 1909 TGGGCCATTCAAATGGAAAAAAGCTGCAAAAAAGGAAATCGCAAAAGAACGTGTTGT 1968
 QY 601 AlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMetIle 620
 Db 1969 GCAGAACATTTTGAGGAAGTACAATGAGGCCCTACAAATTAATGACACAAATTCGAATGATA 2028
 QY 621 AspAlaTyrThrHisLeuGluThrPheTyrAsnGluGluLysAspLysLysPheAlaVal 640
 Db 2029 GATGCGTATACTCATCTTGAAACTTTCTATAATGAAGAGAAAGATAAGAAAGTTTGCAGTC 2088
 QY 641 IleGluAspAspSerAspGluGlyGlyAspAspGluTyrCysAspGlyAspGluAspGlu 660
 Db 2089 ATAGAAGATGATAGTAGGAGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2148
 QY 661 AspAspLeuLysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeuPhe 680
 Db 2149 GATGATTTAAAGAAACCTTTGAAACTGGATGAAACAGATAGATTTCTCATGACTTTATTT 2208
 QY 681 PheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGluLys 700
 Db 2209 TTTGAAAAACAATAAAATGTTGAAAAGGCTGGCTGAAAACCCAGAAATATGAAAATGAAA 2268
 QY 701 LeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArgThrGluGluSerAlaArg 720
 Db 2269 CTGACCAAAATTAAGAAATACCATAATGGAGCAATATACTAGGACTGAGGAATCAGCACGA 2328
 QY 721 GlyIleIlePheThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTrpIleThrGlu 740
 Db 2329 GGAATAATCTTTACAAAAACACGACAGAGTGCATATGCGCTTTCCAGTGGATTACTGAA 2388
 QY 741 AsnGluLysPheAlaGluValGlyValLysAlaHisHisLeuIleGlyAlaGlyHisSer 760
 Db 2389 AATGAAAAATTTGCTGAAGTAGGAGTCAAGGCCCAACCATCTGATTTGGAGTTGGACACAGC 2448
 QY 761 SerGluPheLysProMetThrGlnAsnGluGlnLysGluValIleSerLysPheArgThr 780

```

Db      2449  ||||| ACTGAGTTCAAACCCATGACACAGAATGAACAAAAAGAAGTCATTAGTAAATTTCGCAC 2508
Qy      781   GlyLysIleAsnLeuLeuIleAlaThrThrValAlaGluGluGlyLeuAspIleLysGlu 800
Db      2509  GGAATAATCAATCTGTTATCGTACCACAGTGGCAGAGAAGGCTTGGATATTAAAGAA 2568
Qy      801   CysAsnIleValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAlaArg 820
Db      2569  TGTAAACATTGTTATCCGTTATGGTCTCGTCACCAATGAAATAGCCATGGTCCAGGCCCGT 2628
Qy      821   GlyArgAlaArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGlyVal 840
Db      2629  GGTGAGCCAGAGCTGATGAGAGCACCTACGTCCTGGTTGCTCACAGTGGTTCAGGAGTT 2688
Qy      841   IleGluHisGluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHisCys 860
Db      2689  ATCGAACATGAGACAGTAAATGATTTCCGAGAGAAGATGATGTATAAAGCTATACATTGT 2748
Qy      861   ValGlnAsnMetLysProGluGluTyrAlaHisLysIleLeuGluLeuGlnMetGlnSer 880
Db      2749  GTTCAAAATATGAACACAGAGGAGTATGCTCATAAGATTTTGGAAATTACAGATGCAAGT 2808
Qy      881   IleMetGluLysLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAsnPro 900
Db      2809  ATAATGGAAGAAGAAATGAAACCAAGAGAAATATTGCCAAGCATTACAAGAAATAACCCA 2868
Qy      901   SerLeuIleThrPheLeuCysLysAsnCysSerValLeuAlaCysSerGlyGluAspIle 920
Db      2869  TCACTAATAACTTTCTTGGCAAAACCTGCAGTGTGTAGCCTGTTCTGGGAAGATATC 2928
Qy      921   HisValIleGluLysMetHisHisValAsnMetThrProGluPhelysGluLeuTyrIle 940
Db      2929  CATGTAATTGAGAAAATGTCATCACGTCAATATGACCCAGAAATTCAGGAACCTTTACATT 2988
Qy      941   ValArgGluAsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluIle 960
Db      2989  GTAAGAGAAAAACAAAGCACTGCAAAAGAGTGTGCCGACTATCAAAATAAATGGTGAATC 3048
Qy      961   IleCysLysCysGlyGlnAlaTyrGlyThrMetMetValHisLysGlyLeuAspLeuPro 980
Db      3049  ATCTGCAAAATGTGGCCAGGCTTGGGAACAATGATGGTGCACAAAGGCTTAGATTTGCCT 3108
Qy      981   CysLeuLysIleArgAsnPheValValPheLysAsnAsnSerThrLysLysGlnTyr 1000
Db      3109  TGTCTCAAAATAAGGAATTTGTAGTGGTTTCAAAAATAATTCAACAAAGAAACAATAC 3168
Qy      1001  LysLysTrpValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCysLeu 1020
Db      3169  AAAAAGTGGGTAGAAATACCTATACATTTCCTCAATCTTGACTATTACAGAAATGCTGTTTA 3228
Qy      1021  PheSerAspGluAsp 1025
Db      3229  TTTAGTGATGAGGAT 3243

RESULT 3
AX300832
LOCUS
DEFINITION Sequence 1 from Patent WO0185955.
ACCESSION AX300832
VERSION AX300832.1 GI:17382110
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
AUTHORS Bahr,G., Cocude,C. and Capron,A.
TITLE Rh16 polypeptides and its fragments and polynucleotides encoding
said polypeptides and therapeutic uses
JOURNAL Patent: WO 0185955-A 1 15-NOV-2001;
Istac (FR) ; INSTITUT PASTEUR DE LILLE (FR)
FEATURES
Location/Qualifiers
```

source

CDS

```

1. .3372
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
155. .3232
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD13049.1"
/db_xref="GI:17382111"
/translation="MSGYSTDENFRYLISCFRARKVMYIQVEPVLDTLFLPAEVKE
QIQRVATSGNQMAVELLLSTLEKGVHLLGWTRFVEALRRTGSPLAARYMNPETLDL
PSPSFENAHDEYQLNLNQPTLVLDKLLVRDVLKMEEEELLTIEDNRRIAAENNGN
ESGVRELLKRIIVOKENWESAFNLVLRQTGNNELVQELTGSDCSESNAEIENLSQVDGP
QVEQLLSTTVQPNLEKEVMENNSSSESSFADSVVSESDTSLAEGSVSCLDSESLGH
NSNMGSDSGTMGSDSDENVAARASPEPELQLRPYQMEVAQPALEGKNIIICLPTGSG
KTRVAVYIAKDHLDDKKKASEPGKIVLVNKLVLVEQLFRKEFQPFLLKMYRVIGLSG
DTQLKISFPEVVKSCDIIISTAQIENSLNLNENGEDAGVQLSDFSFIIDCHHTNK
EAVYNIMRHYLMQKLKNRLKKENKPIPLPQILGLTASPGVGATKQAKAEHILK
LCANLDAFTIKTVKENLDQKNQIQEPCCKFAIADATREDPFKEKLEIMTRIQTQCO
MSPMSDFGTQPYEQWAIQMEKKAKEGNRKESVCAEHLRKYNKALQINDTIRMIDAYT
HLETFYNEEKOKKFAVIDSDSDEGDDDEYCDGDEDDDLKKPLKLDDETRMLTLPFE
NNKMLKRLAENPEYENELTKLRNTIMEQYTRTESARGIIFTKTRQSAYALSQWITE
NEKFAEVGVKAHLIGAGHSSEFPKPMTONQEQVTSKPRFTGKINLLIATTVABEGLDI
KECNIVIRYGLVTNEIAMVQARGARADESTYVLVAHSGSGVIERETVNDFREKMYK
AHCVCQNMKPEEYAHKILELQMSIMEKMKTKRNIKNIKNIKNIKNIKNIKNIKNIK
CSGEDIHVIEKMHHVNMTPFEKELYIVRENKALQKCADYQINGEIIICKGQAWGTM
VHKGLDLPCLKIRNFVVVFKNNSTKKQKKWVELPITFPNLDYSECCLFSDED"
```

ORIGIN

Alignment Scores:
Pred. No.: 0 Length: 3372
Score: 5285.00 Matches: 1020
Percent Similarity: 99.71% Conservative: 2
Best Local Similarity: 99.51% Mismatches: 3
Query Match: 99.51% Indels: 0
DB: 6 Gaps: 0
US-09-515-363C-2 (1-1025) x AX300832 (1-3372)

```

Qy      1 MetSerAsnGlyTyrSerThrAspGluAsnPheArgTyrLeuIleSerCysPheArgAla 20
Db      155  ATGTCGAATGGGTATTCACACAGACGAGAAATTTCCGCTATCTCATCTCGTGCTTCAGGGCC 214
Qy      21 ArgValLysMetTyrIleGlnValGluProValLeuAspTyrLeuThrPheLeuProAla 40
Db      215  AGGTTGAAATGTACATCCAGGTGGAGCCTGTGTGGACTACCTGACCTTTTCTGCCTGCA 274
Qy      41 GluValLysGluGlnIleGlnArgThrValAlaThrSerGlyAsnMetGlnAlaValGlu 60
Db      275  GAGGTGAAGGAGCAGATTTCAGAGGACAGTCGCCACCTCCGGGAAACATGCAGGCAGTTGAA 334
Qy      61 LeuLeuLeuSerThrLeuGluLysGlyValTyrHisLeuGlyTyrThrArgGluPheVal 80
Db      335  CTGCTGCTGAGCACCTTGGAGAAGGAGTCTGGCACCTTGGTGGACTCGGGAATTCGTG 394
Qy      81 GluAlaLeuArgArgThrGlySerProLeuAlaAlaArgTyrMetAsnProGluLeuThr 100
Db      395  GAGGCCCTCCGAGAACCCGGCAGCCCTCTGGCCGCCCGCTACATGAACCTGAGCTCACG 454
Qy      101 AspLeuProSerProSerPheGluAsnAlaHisAspGluTyrLeuGlnLeuLeuAsnLeu 120
Db      455  GACTTGCCCTCTCCATCGTTTGAGAACGCTCATGATGAATATCTCCAACCTGCTGAACCTC 514
Qy      121 LeuGlnProThrLeuValAspLysLeuLeuValArgAspValLeuAspLysCysMetGlu 140
Db      515  CTTTCAGCCCACTCTGGTGGACAAGCTTCTAGTTAGAGACGCTCTGGATAAGTGCATGGAG 574
Qy      141 GluGluLeuLeuThrIleGluAspArgAsnArgIleAlaAlaGluAsnAsnGlyAen 160
Db      575  GAGGAACCTGTTGACAATTTGAAGACAGAAACCGGATTGCTGCTGCAGAAACAATGGAAT 634
Qy      161 GluSerGlyValArgGluLeuLeuLysArgIleValGlnLysGluAsnTrpPheSerAla 180
```

Db 635 GAATCAGGTGTAAGAGAGCTACTAAAAAGGATTGTGCAGAAAGAAAACTGGTTCTCTGCA 694

Qy 181 PheLeuAsnValLeuArgGlnThrGlyAsnAsnGluLeuValGlnGluLeuThrGlySer 200

Db 695 TTTCTGAATGTTCTTCGTCAAACAGGAAACAATGAACCTTGTCCAAGAGTTAAACAGGCTCT 754

Qy 201 AspCysSerGluSerAsnAlaGluIleGluAsnLeuSerGlnValAspGlyProGlnVal 220

Db 755 GATTGTCTGAGAAAGCAATGCAGAGATTGAGAAATTTATCACAAGTTGATGGTCTCAAGTG 814

Qy 221 GluGluGlnLeuLeuSerThrThrValGlnProAsnLeuGluLysGluValTrpGlyMet 240

Db 815 GAAGAGCAACTTCTTTCAACCACAGTTCAGCCAAATCTGGAGAAGGAGGTCTGGGGCATG 874

Qy 241 GluAsnAsnSerSerGluSerSerPheAlaAspSerSerValValSerGluSerAspThr 260

Db 875 GAGAATAACTCATCAGAAATCATCTTTGCAGATTCCTCTGTAGTTTCAGAAATCAGACACA 934

Qy 261 SerLeuAlaGluGlySerValSerCysLeuAspGluSerLeuGlyHisAsnSerAsnMet 280

Db 935 AGTTTGGCAGAAAGAGTGTCAGCTGCTTAGATGAAAGTCTGGACATAACAGCAACATG 994

Qy 281 GlySerAspSerGlyThrMetGlySerAspSerAspGluGluAsnValAlaAlaArgAla 300

Db 995 GGCAGTGATTACGCCACCATGGGAAGTGATTCAGATGAAGAGAATGTGGCAGCAAGAGCA 1054

Qy 301 SerProGluProGluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaLeu 320

Db 1055 TCCCGGAGCCGAAACTCCAGCTCAGGCTTACCAATGGAAGTTGCCAGCCAGCCTTG 1114

Qy 321 GluGlyLysAsnIleIleIleCysLeuProThrGlySerGlyLysThrArgValAlaVal 340

Db 1115 GAAGGGAAGAAATATCATCATCTGCCTCCCTACAGGGAGTGGAAAAACCAGAGTGGCTGTT 1174

Qy 341 TyrIleAlaLysAspHisLeuAspLysLysLysAlaSerGluProGlyLysValIle 360

Db 1175 TACATTGCCAAGGATCACTTAGACAAGAGAAAAAGCATCTGAGCCTGGAAAAAGTTATA 1234

Qy 361 ValLeuValAsnLysValLeuLeuValGluGlnLeuPheArgLysGluPheGlnProphe 380

Db 1235 GTTCTTGTCAATAAGGTACTGCTAGTTGAACAGCTCTTCCGCAAGGAGTTCCACCATTT 1294

Qy 381 LeuLysLysTrpTyrArgValIleGlyLeuSerGlyAspThrGlnLeuLysIleSerPhe 400

Db 1295 TTGAAGAAATGGTATCGTGTATTGGATTAAAGTGGTGATACCCAACTGAAAAATATCATTT 1354

Qy 401 ProGluValValLysSerCysAspIleIleSerThrAlaGlnIleLeuGluAsnSer 420

Db 1355 CCAGAAAGTTGTCAAGTCCTGTGATATTATTCAGTACAGCTCAAATCCTTGAAAACTCC 1414

Qy 421 LeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPheSerLeuIle 440

Db 1415 CTCTTAAACTTGAAAAATGGAGAAGATGCTGGTGTTCAAATTGTCAGACTTTTCCCTTCATT 1474

Qy 441 IleIleAspGluCysHisHisThrAsnLysGluAlaValTyrAsnAsnIleMetArgHis 460

Db 1475 ATCATTGATGAATGTCATCACACCAACAAGAGCAGTGTATATAACATCATGAGGCAT 1534

Qy 461 TyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysLysGluAsnLysProValIlePro 480

Db 1535 TATTTGATGCAGAAAGTTGAAAAACAATAGACTCAAGAAAGAAAAACAACCAAGTGTATCCC 1594

Qy 481 LeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyGlyAlaThrLysGlnAla 500

Db 1595 CTTCTCAGATACCTGGGACTAACAGCTTCACCTGGTGTGGAGGGGCCACGAAGCAAGCC 1654

Qy 501 LysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThrIleLysThr 520

Db 1655 AAAGCTGAAGAACACATTTTAAAACTATGTGCCAATCTTGTGATGCAATTTACTATTAATACT 1714

Qy 521 ValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluProCysLysLysPheAla 540

Db 1715 GTTAAAGAAAAACCTTGATCAACTGAAAAACCAAAATACAGGAGCCATGCAAGAAAGTTTGCC 1774

Qy 541 IleAlaAspAlaThrArgGluAspProPheLysGluLysLeuLeuGluIleMetThrArg 560

Db 1775 ATTGCAGATGCAACCAGAGAAGATCCATTAAAGAGAAAACTTTAGAAAAATAATGACAAG 1834

Qy 561 IleGlnThrTyrCysGlnMetSerProMetSerAspPheGlyThrGlnProTyrGluGln 580

Db 1835 ATTCAAACTTATTGTCAAATGAGTCCAATGTCCAGATTTTGGAACTCAACCCCTATGAACAA 1894

Qy 581 TrpAlaIleGlnMetGluLysLysAlaAlaLysLysGlyAsnArgLysGluArgValCys 600

Db 1895 TGGGCCATTCAAATGGAATAAAAGCTGCAAAAAGAGGAAATCGCAAGAAAGTGTTTGT 1954

Qy 601 AlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMetIle 620

Db 1955 GCAGAACATTTGAGGAAGTACAATAAGGCCCTACAAATTAATGACACAAATTCGAATGATA 2014

Qy 621 AspAlaTyrThrHisLeuGluThrPheTyrAsnGluGluLysAspLysLysPheAlaVal 640

Db 2015 GATCGGTATACTCATCTTGAACCTTTCTATAATGAAGAGAAAGATAAGAAAGTTTGCAGTC 2074

Qy 641 IleGluAspAspSerAspGluGlyGlyAspAspGluTyrCysAspGlyAspGluAspGlu 660

Db 2075 ATAGAAGATGATAGTATGATGAGGGTGGTGATGATGATGATGATGATGATGATGATGATG 2134

Qy 661 AspAspLeuLysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeuPhe 680

Db 2135 GATGATTTAAAGAAACCTTTGAAACTGGATGAAACAGATAGATTTCTCATGACTTTATTT 2194

Qy 681 PheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGluLys 700

Db 2195 TTTGAAAAACAATAAAATGTTGAAAAGGCTGGCTGAAACCCAGAAATATGAAAAATGAAAAG 2254

Qy 701 LeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArgThrGluGluSerAlaArg 720

Db 2255 CTGACCAAATTAAGAAATACCATAATGGAGCAATATACTAGGACTGAGGAATCAGCACGA 2314

Qy 721 GlyIleIlePheThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTrpIleThrGlu 740

Db 2315 GGAATAATCTTTACAAAAACACGACAGAGTGCAATATGCGCTTTCCCAAGTGGATTACTGAA 2374

Qy 741 AsnGluLysPheAlaGluValGlyValLysAlaHisHisLeuIleGlyAlaGlyHisSer 760

Db 2375 AATGAAAAAATTTGCTGAAGTAGGAGTCAAAGCCCAACCATCTGATTGGAGCTGGACACAGC 2434

Qy 761 SerGluPheLysProMetThrGlnAsnGluGlnLysGluValIleSerLysPheArgThr 780

Db 2435 AGTGAGTTCAAACCCCATGACACAGAAATGAACAAAAAGAAAGTCATTAGTAAATTTTCGCACT 2494

Qy 781 GlyLysIleAsnLeuLeuIleAlaThrThrValAlaGluGluGlyLeuAspIleLysGlu 800

Db 2495 GGAATAATAAATCTGCTTATCGCTACCAACAGTGGCAGAAAGGCTCTGGATATTAAAGAA 2554

Qy 801 CysAsnIleValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAlaArg 820

Db 2555 TGTAACATTTGTTATCCGTTATGGTCTCGTCACCAATGAAATAGCCATGGTCCAGGCCCGT 2614

Qy 821 GlyArgAlaArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGlyVal 840

Db 2615 GGTCCAGCCAGAGCTGATGAGAGCACCTACGTCCTGGTTGCTCACAGTGGTTCAGGAGTT 2674

Qy 841 IleGluHisGluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHisCys 860

Db 2675 ATCGAACGTGAGACAGTAAATGATTTCCGAGAGAAGATGATGTATAAAGCTATACATTGT 2734

Qy 861 ValGlnAsnMetLysProGluGluTyrAlaHisLysIleLeuGluLeuGlnMetGlnSer 880

Db 2735 GTTCAAAATATGAAACCAGAGGAGTATGCTCATAAAGATTTTGGAAATTACAGATGCAAAAGT 2794

Qy 881 IleMetGluLysLysMetLysThrLysArgAsnIleAlaLysHisThrLysAsnAsnPro 900

Db 2795 ATAATGAAAAAGAAATGAAAAACCAAGAGAAATATTGCCAAGCATTTACAAGAAATAACCCA 2854

QY 901 SerLeuIleThrPheLeuCysLysAsnCysSerValLeuAlaCysSerGlyGluAspIle 920
Db 2855 TCACATAAATCTTCCTTGC AAAA AACTGCAGTGTGCTAGCCTGTTCTGGGAAGATATC 2914
QY 921 HisValIleGluLysMetHisHisValAsnMetThrProGluPhelysGluLeuTyrIle 940
Db 2915 CATGTAATTGAGAAAATGCAATCAGTCATATGACCCAGAAATCAAGGAACCTTTACATT 2974
QY 941 ValArgGluAsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluIle 960
Db 2975 GTAAGAGAAAACAAAGCACTGCAAAAGAAAGTGTGCCGACTATCAAAATAAATGGTGAATC 3034
QY 961 IleCysLysCysGlyGlnAlaTirpGlyThrMetMetValHisLysGlyLeuAspLeuPro 980
Db 3035 ATCTGCAAAATGTGCCAGGCTTGGGGAACAATGATGGTGCACAAAGGCTTAGATTGCCT 3094
QY 981 CysLeuLysIleArgAsnPheValValPheLysAsnAsnSerThrLysLysGlnTyr 1000
Db 3095 TGTCTCAAAATAAGGAATTTTGTAGTGGTTTCAAAAATAATTCAACAAAGAAACAATAC 3154
QY 1001 LysLysTrpValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCysLeu 1020
Db 3155 AAAAAGTGGTAGAATTACCTATCACATTTCCCAATCTTGACTATTCAAGATGCTGTTA 3214
QY 1021 PheSerAspGluAsp 1025
Db 3215 TTTAGTGATGAGGAT 3229
RESULT 4
AY017378
LOCUS
DEFINITION Homo sapiens RNA helicase-DEAD box protein RH116 mRNA, complete cds.
ACCESSION AY017378
VERSION AY017378.1 GI:12621065
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3373)
AUTHORS Cocude,C., Truong,M.-J., Billaut-Mulot,O., Delsart,V., Darcissac,E., Capron,A., Mouton,Y. and Bahr,G.M.
TITLE A novel cellular RNA helicase, RH116, differentially regulates cell growth, programmed cell death and human immunodeficiency virus type 1 replication
JOURNAL J. Gen. Virol. 84 (12), 3215-3225 (2003)
REFERENCE 2 (bases 1 to 3373)
AUTHORS Cocude,C., Kolesnitchenko,V., Billaut-Mulot,O., Truong,M.-J., Capron,A. and Bahr,G.M.
TITLE Identification of a new RNA helicase (RH116) regulated by the immunomodulator Murabutide
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 3373)
AUTHORS Cocude,C., Kolesnitchenko,V., Billaut-Mulot,O., Truong,M.-J., Capron,A. and Bahr,G.M.
TITLE Direct Submission
JOURNAL Submitted (09-JAN-2001) Laboratoire d'Immunologie Moleculaire de l'Infection et de l'Inflammation, Institut Pasteur de Lille, 1 rue du Professeur Calmette BP 245, Lille 59019, France
FEATURES
source
1. .3373
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
155. .3232
/codon_start=1
/product="RNA helicase-DEAD box protein RH116"
/protein_id="AAG54076.1"
/db_xref="GI:12621066"
CDs
155. .3232
/translation="MSGYSTDENFRYLISCFRARVKMYIQVEPVLDTFLPAEYKE QIORTVATSGNMQAVELLSLEKGVWHLGWTRFVRLRRTGSPLAARYMNPFLTDL PSPSFENAHDEYQLQLNLLQPTLVLDKLLVRDVLDKMBEELLTIEDRNRITAAENNGN

ESGVRELLKRIIVQKENWFSAFNLNVLRTQNNELVQELTGSDCSESNAEIENLSQVDGP
QVEEQLLSTTVOPNLEKEVGMENNSSSESFADSSVVSSESDTSLAEGSVSCLDESIGH
NSNMGSDSGTMGSDSDEENVAARASPEPELQLRPYQMEVAQPALEGKNIICLPTGSG
KTRVAVYIAKDHLDKKKASEPKVIVLVNKVLLVEQLFRKEFPFLKKWRYRVIGLSG
DTOLKISFPEVVKSCDIITSTAQILENSLLNLENGEDAGVQLSDFSFIIDECHHTNK
EAVYNNIMRHYLMQKLKNNRLKKENKPVIPLPQILGLTASPGVGATQKAKAEHILK
LCANLDAFTIKTVKENLDQKNQIQEPCCKFAIADATREDPFKEKLLLEIMTRIQTCCQ
MSPMSDFGTQPYEQWAIQMEKKAKEGNRKESVCAEHLRKYNKALQINDTIRMIDAYT
HLETFYNEEKDKKFAVIEDSDDEGDDDEYCDGDEDDDLKKPLKLDETDRFLMTLPFE
NNKMLKRLAENPEYENEKLTKLNTIMEQYTRTESARGIIFTKTRQSAYALSQWITE
NEKFAEVGVKAHHLIGAGHSSEFPKPMQTONEQKEVISKERTGKINLLIATTTVAEGLDI
KECNIVIRYGLVTNEIAMVQARGARADESTYVLVAHSGSGVIERETVNDFRKMYK
AIHCYQNMKPEEYAHKILELQMSIMEKKMTKKNIAKHYNKWNPSLITFLCKNCVLA
CSGEDIHVIEKMHVNMTPEFKELYIVRENKALOKKADYQINGEIIICKGQAWGTM
VHKGLDLPCLKIRNFVVVKNNSTKKQYKKWVELPITFPNLDYSECCLFSDED"
1148. .1201
/note="Region: DEAD-like helicase domain"
1481. .1627
/note="Region: DEAD/DEAH box"
2606. .2647
/note="conserved C-terminal helicase region; Region: helicase C"
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 3373
Score: 5285.00 Matches: 1020
Percent Similarity: 99.71% Conservative: 2
Best Local Similarity: 99.51% Mismatches: 3
Query Match: 99.51% Indels: 0
DB: 9 Gaps: 0
US-09-515-363C-2 (1-1025) x AY017378 (1-3373)
QY 1 MetSerAsnGlyTyrSerThrAspGluAsnPheArgTyrLeuIleSerCysPheArgAla 20
Db 155 ATGTCGAATGGGTATTCACAGACGAGAATTTCCGCTATCTCATCTCGTGTTCAGGGCC 214
QY 21 ArgValLysMetTyrIleGlnValGluProValLeuAspTyrLeuThrPheLeuProAla 40
Db 215 AGGGTGAATAATGTACATCCAGGTGGAGCCTGTGTGGACTACCTGACCTTTCTGCCTGCA 274
QY 41 GluValLysGluGlnIleGlnArgThrValAlaThrSerGlyAsnMetGlnAlaValGlu 60
Db 275 GAGGTGAAGGAGCAGATTTCAGAGGACAGTCGCCACCTCCGGGAACATGCAGGCAGTTGAA 334
QY 61 LeuLeuLeuSerThrLeuGluLysGlyValTrpHisLeuGlyTrpThrArgGluPheVal 80
Db 335 CTGCTGCTGAGCACCTTGGAGAAGGGAGTCTGGCACCTTGGTTGGACTCGGGAATTCGTG 394
QY 81 GluAlaLeuArgArgThrGlySerProLeuAlaAlaArgTyrMetAsnProGluLeuThr 100
Db 395 GAGGCCCTCCGGAGAACCGGCAGCCCTCTGGCCGCCCTACATGAACCTGAGCTCAG 454
QY 101 AspLeuProSerProSerPheGluAsnAlaHisAspGluTyrLeuGlnLeuLeuAsnLeu 120
Db 455 GACTTGCCCTCTCCATCGTTTGAGAACGCTCATGATGAATATCTCCAACCTGCTGAACCTC 514
QY 121 LeuGlnProThrLeuValAspLysLeuLeuValArgAspValLeuAspLysCysMetGlu 140
Db 515 CTTGAGCCCACTCTGGTGGACAAGCTTCTAGTTAGAGACGCTTGGATAAGTGCAATGGAG 574
QY 141 GluGluLeuLeuThrIleGluAspArgAsnArgIleAlaAlaGluAsnAsnGlyAsn 160
Db 575 GAGGAACCTGTGACAATTGAAGACAGAACCCGGATTGCTGCTGCAGAAAAACAATGGAAT 634
QY 161 GluSerGlyValArgGluLeuLeuLysArgIleValGlnLysGluAsnTrpPheSerAla 180
Db 635 GAATCAGGTGTAAGAGAGAGCTACTAAAAAGGATTGTGCAGAAAGAAAACTGGTTCTCTGCA 694
QY 181 PheLeuAsnValLeuArgGlnThrGlyAsnAsnGluLeuValGlnGluLeuThrGlySer 200
Db 695 TTTCTGAATGTTCTTCGTCAAACAGGAAAAACAATGAACTTGTCCAAGAGTTTAACAGGCTCT 754

Db 841 GGCAGGGATTcAGGCACCATGGGAAGTGATTcAGATGAAAGTGTCATCCAGACAAAAAGA 900
 QY 300 AlaSerProGluProGluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAla 319
 Db 901 GTATCCCCGAGCCAGAACTGcAGCTCAGGCCTTACCAATGGAAGTGGCCCAACCAGCT 960
 QY 320 LeuGluGlyLysAsnIleIleCysLeuProThrGlySerGlyLysThrArgValAla 339
 Db 961 CTAGATGGGAAGAATATTATTATCTGCCTCCCCACGGGAGTGGGAAACCAGAGTGGCT 1020
 QY 340 ValTyrIleAlaLysAspHisLeuAspLysLysLysLysAlaSerGluProGlyLysVal 359
 Db 1021 GTTTACATCACCAAGATCACTTAGACAAGAGAAGCAGGCATCTGAATCCGGGAAGGTT 1080
 QY 360 IleValLeuValAsnLysValLeuLeuValGluGlnLeuPheArgLysGluPheGlnPro 379
 Db 1081 ATCGTTCTTGTCAAATAAGGTAATGTGTAGCAGAACAACTTTCCGAAAAGAGTTCAACCCA 1140
 QY 380 PheLeuLysLysTrpTyrArgValIleGlyLeuSerGlyAspThrGlnLeuLysIleSer 399
 Db 1141 TATTTGAAGAAATGGTATCGAATTATTGGATTAAAGTGGCGATACCCAGCTGAAAATATCA 1200
 QY 400 PheProGluValValLysSerCysAspIleIleIleSerThrAlaGlnIleLeuGluAsn 419
 Db 1201 TTTCCAGAAGTTGTCAAATCTTTACGATGTTATTATCAGCACTGCTCAAATCCTTGAAAC 1260
 QY 420 SerLeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPheSerLeu 439
 Db 1261 TCCCTCTTAAATCTGGAGAGTGGAGACGATGACGGTGTGCAGCTGTcAGACTTCTCTCTC 1320
 QY 440 IleIleIleAspGluCysHisHisThrAsnLysGluAlaValTyrAsnAsnIleMetArg 459
 Db 1321 ATTATCATTGATGAGTGCCATCACACCAACAAAGGAGGCGAGTCTATAACAACATCATGAGA 1380
 QY 460 HisTyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysLysGluAsnLysProValIle 479
 Db 1381 CGATATTTGAAGCAGAAGCTGAGAAAACAATGACCTCAAGAAAACAAAACCAACCAGCCATT 1440
 QY 480 ProLeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyGlyAlaThrLysGln 499
 Db 1441 CCCCTGCCGCAGATACTAGGACTGACAGCTTCACCTGGTGTGGAGCAGCCAAAAGCAG 1500
 QY 500 AlaLysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThrIleLys 519
 Db 1501 TCTGAGGCTGAAAAACATATTTTAAATATATGTGCCAATCTTGATGCCTTTACCATTAA 1560
 QY 520 ThrValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluProCysLysLysPhe 539
 Db 1561 ACAGTGAAAGAGAATCTTGGTCAACTCAAAACACCAATAAAGGAACCATGCAAGAAATTT 1620
 QY 540 AlaIleAlaAspAlaThrArgGluAspProPheLysGluLysLeuLeuGluIleMetThr 559
 Db 1621 GTGATTGTGTGACACcAGAGAAAATCCATTAAAGAGAACTTCTAGAAATTATGGCA 1680
 QY 560 ArgIleGlnThrTyrCysGlnMetSerProMetSerAspPheGlyThrGlnProTyrGlu 579
 Db 1681 AGCATTcAGACTTACTGCCAAAAAGTCCAATGTcAGATTTTGGAAACCCCAcATTATGAG 1740
 QY 580 GlnTrpAlaIleGlnMetGluLysLysAlaAlaLysLysGlyAsnArgLysGluArgVal 599
 Db 1741 CAGTGGGCCATTCAAAATGGAGAAAAAGTGTAAAGACGGAAATCGCAAGATCGCGTC 1800
 QY 600 CysAlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMet 619
 Db 1801 TGTGCAGAACATTGTAGGAAGTACAAcGAAGCCCTACAAATCAACGACACGATCCGAATG 1860
 QY 620 IleAspAlaTyrThrHisLeuGluThrPheTyrAsnGluGluLysAspLysLysPheAla 639
 Db 1861 ATTGATGCATATACCCACCTGGAGACATTCTACACTGATGAGAAAAGAAAGAGTTCGCA 1920
 QY 640 ValIleGluAspAspSerAspGluGlyGlyAspAspGluTyrCysAspGlyAspGluAsp 659
 Db 1921 GTCCTC---AATGACAGCGACAAGAGTGTAGACGAGGCCAGCGAGTTGCAATGACCAACTT 1977

QY 660 GluAspAspLeuLysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeu 679
 Db 1978 AAGGGCGATGTAAAGAAATCTTTGAAACTGGACGAAACGGATGAATTTCTCATGAATTTG 2037
 QY 680 PhePheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGlu 699
 Db 2038 TTCTTTGATAACAAGAAAATGTTGAAAAAACTAGCTGAAAAACCCAAAATACGAGAATGAA 2097
 QY 700 LysLeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArgThrGluGluSerAla 719
 Db 2098 AAACCTCATTAATTAAGAAAACACGATACTACTGAACAATTCACAAGCTCTGAGGAGTCCTCC 2157
 QY 720 ArgGlyIleIlePheThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTrpIleThr 739
 Db 2158 CGAGGAATTATTTTACAAAACACGACAGAGCACCTACGCACTTTCCAGTGGATCATG 2217
 QY 740 GluAsnGluLysPheAlaGluValGlyValLysAlaHisIleLeuIleGlyAlaGlyHis 759
 Db 2218 GAAAATGCAAAAGTTTGGGAAGTTGGAGTCAAAAGCGCATCACCTGATTGGCGGGGCAC 2277
 QY 760 SerSerGluPheLysProMetThrGlnAsnGluGlnLysGluValIleSerLysPheArg 779
 Db 2278 AGCAGTGAAGTCAAGCCCATGACTCAGACTGAAACAAAAGAAAGTCATTAGTAAATTTCCG 2337
 QY 780 ThrGlyLysIleAsnLeuLeuIleAlaThrThrValAlaGluGluGlyLeuAspIleLys 799
 Db 2338 ACTGGGAAATAAATCTGCTTATCGCTACGCGTGGCAGGAGGAGGAGGCTGGATATCAA 2397
 QY 800 GluCysAsnIleValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAla 819
 Db 2398 GAGTGAATATTGTTATTGCTTATGGCCTTGTcACGAACGAGATAGCCATGGTCCAGGCC 2457
 QY 820 ArgGlyArgAlaArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGly 839
 Db 2458 CGGGTcGAGCCAGAGCTGATGAAAGCACGCTATGTCTGCTCACCAGCAGTGGTcCAGGA 2517
 QY 840 ValIleGluHisGluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHis 859
 Db 2518 GTTACCGAACGGGAGATTGTTAATGATTTCCGAGAGAAGATGATGATATAAAGCTATTAAc 2577
 QY 860 CysValGlnAsnMetLysProGluGluTyrAlaHisLysIleLeuGluLeuGlnMetGln 879
 Db 2578 CGTGTTCAAAACATGAAACcAGAGGAGTATGCACATAAGATTTTGGAAATTGCAGGTGCAA 2637
 QY 880 SerIleMetGluLysLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAsn 899
 Db 2638 AGTATCTTGAAAAAGAAAATGAAAGTCAAAAGAAAGCATTTGCAAGCAATACACGACAAT 2697
 QY 900 ProSerLeuIleThrPheLeuCysLysAsnCysSerValLeuAlaCysSerGlyGluAsp 919
 Db 2698 CCATCGTTAATAACACTTCTCTGCAAAAATTTAGCATGTGTGCTGCTCGGGAGAAAAc 2757
 QY 920 IleHisValIleGluLysMetHisHisValAsnMetThrProGluPheLysGluLeuTyr 939
 Db 2758 ATCCATGTcATTGAGAAGATGCATCATGTCAATATGACACCAGAAATTCAAGGGACTCTAC 2817
 QY 940 IleValArgGluAsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGlu 959
 Db 2818 ATTGTAAGAGAAAAcAAAGCACTGCAAAAGAAATTTGCTGATTATCAGACCAATTGGAGAG 2877
 QY 960 IleIleCysLysCysGlyGlnAlaTrpGlyThrMetMetValHisLysGlyLeuAspLeu 979
 Db 2878 ATTATCTGCAAGTGTGGCCAGGCTTGGGGAACAAATGATGGTGCACAAAGGTTTAGATTG 2937
 QY 980 ProCysLeuLysIleArgAsnPheValValPheLysAsnAsnSerThrLysLysGln 999
 Db 2938 CCTGTCTTAAAAATAAGGAATTTTGTAGTCAATTTCAAAAATAACTCACCGAAGAAACAG 2997
 QY 1000 TyrLysLysTrpValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCysCys 1019
 Db 2998 TACAAGAAGTGGGTGGAAATTGCCTATCAGATTTCTCTGATCTTGTACTACTCAGAAATACTGC 3057

Db 2035 TGTGCAGAACATTTGAGGAAGTACACGAAGCCCTACAATAACGACACGATCCGAATG 2094
QY 620 IleAspAlaTyrThrHisLeuGluThrPheTyrAsnGluGluLysAspLysLysPheAla 639
Db 2095 ATTGATGCATATAGCCACCTGGAGGCATTTACACTGATGAGAGAAAGAAAGAGTTGCGCA 2154
QY 640 ValIleGluAspAspSerAspGluGlyGlyAspAspGluTyrCysAspGlyAspGluAsp 659
Db 2155 GTCCTC--AATGACAGCGACAAGAGTGATGACGAGGCCAGCAGTTGCAATGACCAACTT 2211
QY 660 GluAspAspLeuLysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeu 679
Db 2212 AAGGGCGATGTAAGAAATCTTTGAAACTGGACGAAACGGATGAATTTCTCATGAATTG 2271
QY 680 PhePheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGlu 699
Db 2272 TTCTTTTGATAACAAGAAATGTTGAAAAACTAGCTGAAACCCAAATACGAGAATGAA 2331
QY 700 LysLeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArgThrGluGluSerAla 719
Db 2332 AAACCTCATTAATAAGAAACACGATACCTGGAACAATTCACAAGGTCTGAGGAGTCCTCC 2391
QY 720 ArgGlyIleIlePheThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTrpIleThr 739
Db 2392 CGAGGAATTAATTTTACAAAAACACGACAGAGCACCTACGCCACTTCCCAGTGGATCATG 2451
QY 740 GluAsnGluLysPheAlaGluValGlyValLysAlaHisHisLeuIleGlyAlaGlyHis 759
Db 2452 GAAATGCAAGTTTGGGAAGTTGGAGTCAAAGCGCATCACCTGATTTGGCGGGGCAC 2511
QY 760 SerSerGluPheLysProMetThrGlnAsnGluGlnLysGluValIleSerLysPheArg 779
Db 2512 AGCAGTGAAGTCAAGCCCATGACTCAGACTGAACAAAAAGAGTCAATTAGTAATTCGC 2571
QY 780 ThrGlyLysIleAsnLeuIleAlaThrThrValAlaGluGluGlyLeuAspIleLys 799
Db 2572 ACTGGCGAAATAATCTGCTTATCGCTACGCGTGGCAGAGGAAGGCCTGGATATCAAA 2631
QY 800 GluCysAsnIleValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAla 819
Db 2632 GAGTGCATATTTGTTATTCGTTATGGCTTGTTCACGAACGAGATAGCCATGGTCCAGGCC 2691
QY 820 ArgGlyArgAlaArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGly 839
Db 2692 CGGGGTCGAGCCAGAGCTGATGAAAGCACGATGTCTGTGTCACCAGCAGTGGCTCAGGA 2751
QY 840 ValIleGluHisGluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHis 859
Db 2752 GTTACCGAACGGGAGATTGTTAATGATTTCCGAGAGAAGATGATGTATAAGCTATTAAAC 2811
QY 860 CysValGlnAsnMetLysProGluGluTyrAlaHisLysIleLeuGluLeuGlnMetGln 879
Db 2812 CGTGTTCAAAACATGAAACCAGAGGAGTATGCACATAAGATTTTGAATTGCAGGTGCAA 2871
QY 880 SerIleMetGluLysLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAsn 899
Db 2872 AGTATCCTGGAAAAGAAAAATGAAAGTCAAAAGAGCATTTGCAACCAATACACGACAA 2931
QY 900 ProSerLeuIleThrPheLeuCysLysAsnCysSerValLeuAlaCysSerGlyGluAsp 919
Db 2932 CCATCGTTAATAACACTTCTCTGCAAAAATTGTAGCATGTCTGCTCGGGAGAAAAC 2991
QY 920 IleHisValIleGluLysMetHisHisValAsnMetThrProGluPheLysGluLeuTyr 939
Db 2992 ATCCATGTCTATTGAGAAGATGCATCATGTCAATATGACACCAGAAATTCGAAGGACTCTAC 3051
QY 940 IleValArgGluAsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGlu 959
Db 3052 ATTGTAAGAGAAAAACAAGCACTGCAAAAGAAATTTGCTGATTATCAGACCAATGGAGAG 3111
QY 960 IleIleCysLysCysGlyGlnAlaTrpGlyThrMetMetValHisLysGlyLeuAspLeu 979

Db 3112 ATTATCTCAAGTGTGGCCAGGCTTGGGAAACAATGATGGTGCAAAAAGGTTTAGATTG 3171
QY 980 ProCysLeuLysIleArgAsnPheValValPheLysAsnAsnSerThrLysLysGln 999
Db 3172 CCTTGTCTTAAATAAGGAATTTTGTAGTCAATTTCAAAAATACTACCGAAGAAACAG 3231
QY 1000 TyrLysLysTrpValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCysCys 1019
Db 3232 TACAAGAAGTGGTGAATTCCTATCAGATTTCTGTATCTTGACTACTCAGAAATACTGC 3291
QY 1020 LeuPheSerAspGluAsp 1025
Db 3292 TTGTATAGTGATGAAGAT 3309
RESULT 8
BC078180
LOCUS
DEFINITION
Homo sapiens interferon induced with helicase C domain 1, mRNA
(CDNA clone IMAGE:6502757), partial cds.
ACCESSION
BC078180
VERSION
BC078180.1 GI:50415794
KEYWORDS
Homo sapiens (human)
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 2540)
AUTHORS
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 2540)
Strausberg,R.
Direct Submission
Submitted (19-JUL-2004) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@nhgri.nih.gov
Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,
Maduro,Q.I., Masello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C.,
McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W.,

Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/BLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 174 Row: p Column: 16
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 27886567.

FEATURES

source
1..2540
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6502757"
/tissue_type="Uterus, leiomyosarcoma"
/clone_lib="NIH MGC_71"
/lab_host="DH10B"
/note="vector: pCMV-SPORT6"
1..>2540
/gene="IFIH1"
/note="synonyms: MDAS, MDA-5, Hlcd"
/db_xref="LocusID:64135"
/db_xref="MIM:606951"
224..>2540
/gene="IFIH1"
/codon_start=1
/product="IFIH1 protein"
/protein_id="AAH78180.1"
/db_xref="GI:50415795"
/db_xref="LocusID:64135"
/db_xref="MIM:606951"
/translation="MSGYSTDENFRYLISCFRARVKMYIQVEPVLDTFLPAEVKE
QIORTVATSGNMQAVELLTSTLEKGVHGLGWTREFVEALRRTGSPLAARYNPELTDL
PSPSFENAHDEYLQLLNLLQPTLVLDKLLVRDVLDKCMBEELLTIEDRNRIRIAAENNGN
ESGVRELLKRIQKENWFSAFNLVLRQTGNNELVQELTGSDCSESNAEINLSQVDGP
OVEEQLLSTVPNLEKEVWGMENNSESSESFADSSVVSSEDTSLAEGSVSCLDSESLGH
NSNMGSDSGTWGSDSDEENVARASPEPELQLRPYQMEVAQPALEGKNIICLPTGSG
KTRVAVYIAKDHLDKKKASEPQKIVLVNKKVLLVEQLFRKEFQPFLLKKWYRVIGLSG
DTQLKISFPEVVKSCDIIISTAQILENSLLNLENGEDAGVQLSDFSLIIDECHHTNK
EAVYNNIMRHYLMQKLKNRLKKENKVPVLPQIILGLTASPGVGGATKQAKAEHILK
LCANLDAFTIKTVKENLDQKNQIQEPCKKFAIADATREDPFKEKLEIMTRIQTCCQ
MSPMSDFGTQPYEQWAIQMEKKAKEGNRKERVCAEHLRKYNEALQINDTIRMIDAYT
HLETFYNEKDKKFAVIEDSDGDDDEYCDGEDDEDLKKPLKLDETDRFLMTLFFE
NNKMLKRLAENPEYENEKLTKLNRNTIMEQYTRTEESARGIIFTKTRQSAYALSQWITE
NEKFAEVGVKAHHLIGAGHSSEFKPMTQKKK"

ORIGIN

Alignment Scores:
Pred. No.: 3.45e-259 Length: 2540
Score: 3954.00 Matches: 768
Percent Similarity: 99.87% Conservative: 3
Best Local Similarity: 99.48% Mismatches: 1
Query Match: 74.45% Indels: 0
DB: 9 Gaps: 0
US-09-515-363C-2 (1-1025) x BC078180 (1-2540)

QY	1	MetSerAsnGlyTyrSerThrAspGluAsnPheArgTyrLeuIleSerCysPheArgAla	20
Db	224	ATGTCGAATGGGTATCCACAGACGAGAAATTCGGCTATCTCATCTCGTGTTCAGGGCC	283
QY	21	ArgValLysMetTyrIleGlnValGluProValLeuAspTyrLeuThrPheLeuProAla	40
Db	284	AGGGTGAAAAATGTACATCCAGGTGGAGCCCTGTGCTGGACTACCTTCTGCCTGCA	343
QY	41	GluValLysGluGlnIleGlnArgThrValAlaThrSerGlyAsnMetGlnAlaValGlu	60
Db	344	GAGGTGAAGGAGCAGATTCCAGAGGACAGTCGCCACCTCCGGGAACATGCAGGCAGTTCAA	403
QY	61	LeuLeuLeuSerThrLeuGluLysGlyValTrpHisLeuGlyTrpThrArgGluPheVal	80
Db	404	CTGCTGCTGAGCACCTTGGAGAAGGGAGTCTGGCACCTTGGTTGGACTCGGGAATTCGTG	463

QY	81	GluAlaLeuArgArgThrGlySerProLeuAlaAlaAargTyrMetAsnProGluLeuThr	100
Db	464	GAGGCCCTCCGGAGAAACCGGCAGCCCTCTGGCCGCCGTACATGAACCCCTGAGCTCAG	523
QY	101	AspLeuProSerProSerPheGluAsnAlaHisAspGluTyrLeuGlnLeuLeuAsnLeu	120
Db	524	GACTTGCCCTCTCCATCGTTTGAGAACGCTCATGATGAATATCTCCAACCTGCTGAACCTC	583
QY	121	LeuGlnProThrLeuValAspLysLeuLeuValArgAspValLeuAspLysCysMetGlu	140
Db	584	CTTCAGCCCACTCTGGTGGACAAGCTTCTAGTTAGAGACGCTCTGGATAAGTGCAATGGAG	643
QY	141	GluGluLeuLeuThrIleGluAspArgAsnArgIleAlaAlaGluAsnAsnGlyAsn	160
Db	644	GAGGAACCTGTTGACAATTGAAGACAGAAACCGGATTGCTGTCGAGAAAAACAATGGAAAT	703
QY	161	GlusSerGlyValArgGluLeuLeuLysArgIleValGlnLysGluAsnTrpPheSerAla	180
Db	704	GAATCAGGTGTAAGAGAGAGCTACTAAAAAGGATTGTGCAGAAAGAAAACTGGTTCTCTGCA	763
QY	181	PheLeuAsnValLeuArgGlnThrGlyAsnAsnGluLeuValGlnGluLeuThrGlySer	200
Db	764	TTTCTGAATGTTCTTCGTCAAACAGGAAACAATGAACCTTGTCGAAGAGTTAACAGGCTCT	823
QY	201	AspCysSerGluSerAsnAlaGluIleGluAsnLeuSerGlnValAspGlyProGlnVal	220
Db	824	GATTGCTCAGAAAGCAATGCAGAGATTGAGAAATTTATCACAAAGTTGATGGTCTCAAGTG	883
QY	221	GluGluGlnLeuLeuSerThrThrValGlnProAsnLeuGluLysGluValTrpGlyMet	240
Db	884	GAAGAGCAACTTCTTTCAACCACAGTTCAGCCAAATCTGGAGAAGGAGGTCTGGGGCATG	943
QY	241	GluAsnAsnSerSerGluSerSerPheAlaAspSerSerValValSerGluSerAspThr	260
Db	944	GAGATAACTCATCAGAATCATCTTTTGCAGATTCTTCTGTAGTTTTCAGAAATCAGACACA	1003
QY	261	SerLeuAlaGluGlySerValSerCysLeuAspGluSerLeuGlyHisAsnSerAsnMet	280
Db	1004	AGTTTGGCAGAAGGAAGTGTGAGCTGCTTAGATGAAAGTCTTGACATATAACAGCAACATG	1063
QY	281	GlySerAspSerGlyThrMetGlySerAspSerAspGluGluAsnValAlaAlaArgAla	300
Db	1064	GGCAGTGATTGAGGCACCATGGGAAGTGATTTCAGATGAAGAGAATGTGGCAGCAAGAGCA	1123
QY	301	SerProGluProGluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaLeu	320
Db	1124	TCCCCGGAGCCAGAACTCCAGCTCAGGCCTTACCAAATGGAAAGTTGCCCAGCCCTTG	1183
QY	321	GluGlyLysAsnIleIleIleCysLeuProThrGlySerGlyLysThrArgValAlaVal	340
Db	1184	GAAGGGAAGAAATATCATCATCTGCTCCCTACAGGGAGTGGAAAAACCCAGAGTGGCTGTT	1243
QY	341	TyrIleAlaLysAspHisLeuAspLysLysLysAlaSerGluProGlyLysValIle	360
Db	1244	TACATTGCCAAGGATCACTTAGACAAGAAAAAAGCATCTGAGCCTGGAAAAAGTTATA	1303
QY	361	ValLeuValAsnLysValLeuLeuValGluGlnLeuPheArgLysGluPheGlnProPhe	380
Db	1304	GTTCTTGTCAATAAGGTACTGCTAGTTGAACAGCTCTTCCGCAAGGAGTTCCAACCAATTT	1363
QY	381	LeuLysLysTrpTyrArgValIleGlyLeuSerGlyAspThrGlnLeuLysIleSerPhe	400
Db	1364	TTGAAGAAATGGTATCGTGTATTGGATTAAAGTGGTATACCCCACTGAAAAATATCATTT	1423
QY	401	ProGluValValLysSerCysAspIleIleIleSerThrAlaGlnIleLeuGluAsnSer	420
Db	1424	CCAGAAAGTTGTCAAGTCCTGTGATATTATTATCAGTACAGCTCAAATCCTTGAAAACTCC	1483
QY	421	LeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPheSerLeuIle	440
Db	1484	CTCTTAAACTTGGAAAAATGGAGAAGATGCTGGTGTTCATTTGTTCAGACTTTTCCCTCAT	1543
QY	441	IleIleAspGluCysHisHisThrAsnLysGluAlaValTyrAsnAsnIleMetArgHis	460

QY 221 GluGluGlnLeuLeuSerThrThrValGlnProAsnLeuGluLysGluValTrpGlyMet 240
Db 621 ----- 621
QY 241 GluAsnAsnSerSerGluSerSerPheAlaAspSerSerValValSerGluSerAspThr 260
Db 622 -----GAATCAGACACA 633
QY 261 SerLeuAlaGluGlySerValSerCysLeuAspGluSerLeuGlyHisAsnSerAsnMet 280
Db 634 AGTTTGGCAGAAAGAGTGTAGCTGCTTAGATGAAGCTTTGGACATAACAGCAACATG 693
QY 281 GlySerAspSerGlyThrMetGlySerAspSerAspGluGluAsnValAlaAlaArgAla 300
Db 694 GGCAGTGATTCAGGCACCATGGGAAGTGATTTCAGATGAAGAGAATGTGGCAGCAAGACA 753
QY 301 SerProGluProGluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaLeu 320
Db 754 TCCCGGAGCCAGAACTCCAGCTCAGGCCTTACCAATGGAAGTTGCCAGCCAGCCTTG 813
QY 321 GluGlyLysAsnIleIleIleCysLeuProThrGlySerGlyLysThrArgValAlaVal 340
Db 814 GAAGGGAAGAAATATCATCATCTGCCTCCCTACAGGGAGTGGAAAAACCAGAGTGGCTGT 873
QY 341 TyrIleAlaLysAspHisLeuAspLysLysLysAlaSerGluProGlyLysValIle 360
Db 874 TACATTGCCAAGGATCACTTAGACAAGAAGAAAAAGCATCTGAGCCTGGAAAAAGTTATA 933
QY 361 ValLeuValAsnLysValLeuLeuValGluGlnLeuPheArgLysGluPheGlnProPhe 380
Db 934 GTTCTTGTCATAAGGTACTGCTAGTTGAACAGCTCTTCCGAAGGAGTTCCAACCATTT 993
QY 381 LeuLysLysTrpTyrArgValIleGlyLeuSerGlyAspThrGlnLeuLysIleSerPhe 400
Db 994 TTGAAGAAATGGTATCGTGTATTGGATTAAAGTGGTATACCCAACTGAAAAATATCATTT 1053
QY 401 ProGluValValLysSerCysAspIleIleIleSerThrAlaGlnIleLeuGluAsnSer 420
Db 1054 CCAGAAAGTTGTCAAGTCTGTGATATTATTATCATAGTACAGTCAAAATCCTTGAAAACTCC 1113
QY 421 LeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPheSerLeuIle 440
Db 1114 CTCTTAAACTTGGAAAAATGGAGAAGATCGTGGTGTTCATTGTGCAGACTTTTCCCTCATT 1173
QY 441 IleIleAspGluCysHisHisThrAsnLysGluAlaValTyrAsnAsnIleMetArgHis 460
Db 1174 ATCATTTGATGATGTTCATCACACCAACAAGAGCAGTGTATATAATACATCATGAGCGT 1233
QY 461 TyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysLysGluAsnLysProValIlePro 480
Db 1234 TATTTGATGCAGAAGTTGAAAAACAATAGACTCAAGAAAGAAAAACAACCAAGTATCC 1293
QY 481 LeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyGlyAlaThrLysGlnAla 500
Db 1294 CTTCTCTCAGATACTGGGACTAACAGCTTACCTGGTGTGGAGGGGCCACGAAGAGCC 1353
QY 501 LysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThrIleLysThr 520
Db 1354 AAAGCTGAAGAACACACATTTTAAAACTATGTGCCAATCTTGATGCATTTACTATTAAAACT 1413
QY 521 ValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluProCysLysLysPheAla 540
Db 1414 GTTAAAGAAAAACCTTGATCAACTGAAAAACCAATAACAGGAGCCATGCAAGAGTTTGCC 1473
QY 541 IleAlaAspAlaThrArgGluAspProPheLysGluLysLeuLeuGluIleMetThrArg 560
Db 1474 ATTGCAGATGCACACAGAGAGATCCATTTAAAGAGAACTTCTAGAAATAATGACAAGG 1533
QY 561 IleGlnThrTyrCysGlnMetSerProMetSerAspPheGlyThrGlnProTyrGluGln 580
Db 1534 ATTCAAACTTATGTCAAATGAGTCCAATGTGAGATTTTGGAACTCAACCCCTATGAACAA 1593
QY 581 TrpAlaIleGlnMetGluLysLysAlaAlaLysLysGlyAsnArgLysGluArgValCys 600

Db 1594 TGGGCCATTCAAATGGAAAAAAA----- 1617
QY 601 AlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMetIle 620
Db 1617 ----- 1617
QY 621 AspAlaTyrThrHisLeuGluThrPheTyrAsnGluGluLysAspLysLysPheAlaVal 640
Db 1617 ----- 1617
QY 641 IleGluAspSerAspGluGlyGlyAspAspGluTyrCysAspGlyAspGluAspGlu 660
Db 1617 ----- 1617
QY 661 AspAspLeuLysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeuPhe 680
Db 1617 ----- 1617
QY 681 PheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGluLys 700
Db 1618 ---GAAAAACAATAAATGTTGAAAAGGCTGGCTGAAAAACCCAGAAATATGAAAAATGAAAAG 1674
QY 701 LeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArgThrGluGluSerAlaArg 720
Db 1675 CTGACCAAAATTAGAAATACCATAATGGAGCAATATACTAGGACTGAGGAATCAGCACGA 1734
QY 721 GlyIleIlePheThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTrpIleThrGlu 740
Db 1735 GGAATAATCTTTACAAAAACACGACAGAGTGCATATGCGCTTCCAGTGGATTACTGAA 1794
QY 741 AsnGluLysPheAlaGluValGlyValLysAlaHisLeuIleGlyAlaGlyHisSer 760
Db 1795 AATGAAAAAATTTGCTGAAGTAGGAGTCAAAGCCCCACCATCTGATTGGAGCTGGACACAGC 1854
QY 761 SerGluPheLysProMetThrGlnAsnGluGlnLysGluValIleSerLysPheArgThr 780
Db 1855 AGTGAGTTCAAACCCCATGACACAGAATGAACAAAAAGAGTCATTAGTAAATTCGCACT 1914
QY 781 GlyLysIleAsnLeuLeuIleAlaThrThrValAlaGluGluGlyLeuAspIleLysGlu 800
Db 1915 GGAAAAATAAATCTGCTTATCGTACCACAGTGGCAGAGAAGGTCTGGATATTAAAGAA 1974
QY 801 CysAsnIleValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGln 818
Db 1975 TGTAACATTGTTATCCGTTATGCTCTCGTCACCAATGAAATAGCCATGGTCCAG 2028
RESULT 10
AX714144
LOCUS AX714144 1776 bp DNA linear PAT 15-APR-2003
DEFINITION Sequence 828 from Patent EP1293569.
ACCESSION AX714144
VERSION AX714144.1 GI:29889072
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Isogai,T., Sugiyama,T., Otsuki,T., Wakamatsu,A., Sato,H., Ishii,S., Yamamoto,J.I., Isono,Y., Hio,Y., Otsuka,K., Nagai,K., Irie,R., Tamechika,I., Seki,N., Yoshikawa,T., Otsuka,M., Nagahari,K. and Masuho,Y.
TITLE Full-length cDNAs
JOURNAL Patent: EP 1293569-A 828 19-MAR-2003;
Helix Research Institute (JP) ; Research Association for Biotechnology (JP)
FEATURES
source Location/Qualifiers
1 .1776
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN

Alignment Scores:				
Pred. No.:	5.15e-186	Length:	1776	
Score:	2874.00	Matches:	546	
Percent Similarity:	99.64%	Conservative:	3	
Best Local Similarity:	99.09%	Mismatches:	2	
Query Match:	54.11%	Indels:	0	
DB:	6	Gaps:	0	
US-09-515-363C-2 (1-1025) x AX714144 (1-1776)				
QY	475	AsnLysProValIleProLeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGly	494	
Db	1	CACAAACAGTGATTCCCTTCCTCAGATACTGGGACTAACAGCTTCACCTGGTGTGGA	60	
QY	495	GlyAlaThrLysGlnAlaLysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeuAsp	514	
Db	61	GGGGCCAGCAAGCAAGCCAAAGCTGAAGAACACATTTTAAAACTATGTGCCAATCTTGAT	120	
QY	515	AlaPheThrIleLysThrValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGlu	534	
Db	121	GCATTTACTATTAAACCTGTTAAAGAAAACCTTGATCAACTGAAAAACCAATAACAGGAG	180	
QY	535	ProCysLysLysPheAlaIleAlaAspAlaThrArgGluAspPropheLysGluLysLeu	554	
Db	181	CCATGCAAGAAGTTTGCCATTGCAGATGCAACCAGAGAAGATCCATTTTAAAGAAACTT	240	
QY	555	LeuGluIleMetThrArgIleGlnThrTyrCysGlnMetSerProMetSerAspPheGly	574	
Db	241	CTAGAAATAATGACAAGGATTCAACTTATTGTCAAATGAGTCCAATGTCTAGATTTTGA	300	
QY	575	ThrGlnProTyrGluGlnTrpAlaIleGlnMetGluLysLysAlaAlaLysLysGlyAsn	594	
Db	301	ACTCAACCTATGAACAATGGGCCATTCAAATGGAAAAAAAAGCTGCAAAAGAAAGAAAT	360	
QY	595	ArgLysGluArgValCysAlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsn	614	
Db	361	CGCAAAGAACGTGTTTGTGCAGAACATTTGAGGAAGTACAATGAGGCCCTACAAATTAAT	420	
QY	615	AspThrIleArgMetIleAspAlaTyrThrHisLeuGluThrPheTyrAsnGluGluLys	634	
Db	421	GACACAATTGGAATGATAGATGCGTATACTCATCTTGAAACTTCTATAATGAAGAGANA	480	
QY	635	AspLysLysPheAlaValIleGluAspAspSerAspGluGlyGlyAspAspGluTyrCys	654	
Db	481	GATAAGAAGTTGCAGTCATAGAAGATGATAGTATGAGGGTGGTGATGATGATATTGT	540	
QY	655	AspGlyAspGluAspGluAspLeuLysLysProLeuLysLeuAspGluThrAspArg	674	
Db	541	GATGGTGATGAAGATGAGGATGATTAAAGAAACCTTTGAAACTGGATGAACAGATAGA	600	
QY	675	PheLeuMetThrLeuPhePheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnPro	694	
Db	601	TTTCTCATGACTTTATTTTTTGAAAAACAATAAAATGTTGAAAAGGCTGGCTGAAAAACCCA	660	
QY	695	GluTyrGluAsnGluLysLeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArg	714	
Db	661	GAATATGAAAATGAAAAGCTGACCAAAATTAAGAAATACCAATAATGGAGCAATATACTAGG	720	
QY	715	ThrGluGluSerAlaArgGlyIleIlePheThrLysThrArgGlnSerAlaTyrAlaLeu	734	
Db	721	ACTGAGGAATCAGCACGAGGAATAATCTTTACAAAAACACGACAGAGTGCAATATGCGCTT	780	
QY	735	SerGlnTrpIleThrGluAsnGluLysPheAlaGluValGlyValLysAlaHisLeu	754	
Db	781	TCCCAGTGGATTACTGAAAATGAAAAATTTGCTGAAGTAGAGTCAAAGCCCACTCTG	840	
QY	755	IleGlyAlaGlyHisSerSerGluPheLysProMetThrGlnAsnGluGlnLysGluVal	774	
Db	841	ATTGGAGCTGGACACAGCAGTGAGTGCAAACCCATGACACAGAATGAACAAAAAGATC	900	
QY	775	IleSerLysPheArgThrGlyLysIleAsnLeuLeuIleAlaThrThrValAlaGluGlu	794	

Db	901	ATTAGTAAATTCGCACTGGAAGATAAATCTGCTTATCGCTACCACAGTGGCAGAGAA	960	
QY	795	GlyLeuAspIleLysGluCysAsnIleValIleArgTyrGlyLeuValThrAsnGluIle	814	
Db	961	GGTCTGGATATTAAAGAAATGTAACATTTGTTATCCGTTATGGTCTCGTCACCAATGAATA	1020	
QY	815	AlaMetValGlnAlaArgGlyArgAlaArgAlaAspGluSerThrTyrValLeuValAla	834	
Db	1021	GCCATGGTCCAGGCCCGTGGTCGAGCCAGAGCTGATGAGAGCACCTACGTCTCTGTTGCT	1080	
QY	835	HisSerGlySerGlyValIleGluHisGluThrValAsnAspPheArgGluLysMetMet	854	
Db	1081	CACAGTGGTTCAGGAGTTATCGAACGTCGAGACAGTTAATGATTTCCGAGAGAAGATGATG	1140	
QY	855	TyrLysAlaIleHisCysValGlnAsnMetLysProGluGluTyrAlaHisLysIleLeu	874	
Db	1141	TATAAAGCTATACATTTGTTCAAATATGAAACCCAGAGGAGTATGCTCATAGATTTTG	1200	
QY	875	GluLeuGlnMetGlnSerIleMetGluLysLysMetLysThrLysArgAsnIleAlaLys	894	
Db	1201	GAATTACAGATGCAAAAGTATAATGGAAGAAATAATGAAAACCAAGAGAAATATTGCCAAG	1260	
QY	895	HisTyrLysAsnAsnProSerLeuIleThrPheLeuCysLysAsnCysSerValLeuAla	914	
Db	1261	CATTACAAGAAATAACCCATCCTACTAATAACTTTCCTTTGCAAAAACTGCAGTGTGTAGCC	1320	
QY	915	CysSerGlyGluAspIleHisValIleGluLysMetHisHisValAsnMetThrProGlu	934	
Db	1321	TGTTCTGGGGAAGATATCCATGTAAATTGAGAAAAATGCATCACGTCAATATGACCCAGAA	1380	
QY	935	PheLysGluLeuTyrIleValArgGluAsnLysAlaLeuGlnLysLysCysAlaAspTyr	954	
Db	1381	TTCAAGGAACTTTACATTTGTCCTCAAAATAAGGAATTTTGTAGTGGTTTTCAAAAATAAT	1440	
QY	955	GlnIleAsnGlyGluIleIleCysLysCysGlyGlnAlaTrpGlyThrMetMetValHis	974	
Db	1441	CAAAATAATGGTGAATCATCTGCAAAATGTGGCCAGGCTTGGGAAACAATGATGGTGAC	1500	
QY	975	LysGlyLeuAspLeuProCysLeuLysIleArgAsnPheValValPheLysAsnAsn	994	
Db	1501	AAAGGCTTAGATTTGCCTTGTCTCAAAATAAGGAATTTTGTAGTGGTTTTCAAAAATAAT	1560	
QY	995	SerThrLysLysGlnTyrLysLysTrpValGluLeuProIleThrPheProAsnLeuAsp	1014	
Db	1561	TCAACAAAGAAACAATACAAAAGTGGGTAGAATTACCTATCACATTTCCCAATCTTGAC	1620	
QY	1015	TyrSerGluCysCysLeuPheSerAspGluAsp	1025	
Db	1621	TATTCAGAAATGCTGTTTATTATTAGTAGAGGAT	1653	
RESULT 11				
AK056293				
LOCUS				
DEFINITION				
to Homo sapiens CDNA FLJ31731 fis, clone NT2RI2006855, weakly similar				
to Homo sapiens RNA helicase (RIG-I) mRNA.				
ACCESSION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
1				
Murakawa,K., Fujimori,K., Tanai,H., Kimata,M., Watanabe,M.,				
Kikkawa,E., Omura,Y., Abe,K., Kamihara,K., Katsuta,N., Sato,K.,				
Tanikawa,M., Yamazaki,M., Ninomiya,K., Ishibashi,T., Yamashita,H.,				
Murakawa,K., Suzuki,Y., Nishikawa,T., Otsuki,T., Sugiyama,T., Irie,R.,				
Wakamatsu,A., Hayashi,K., Sato,H., Nagai,K., Kimura,K., Makita,H.,				
Sekine,M., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T.,				
Ishii,S., Yamamoto,J., Saito,K., Kawai,Y., Isono,Y., Nakamura,Y.,				
Nagahari,K., Murakami,K., Yasuda,T., Iwayanagi,T., Wagatsuma,M.,				
Shiratori,A., Sudo,H., Hosoiri,T., Kaku,Y., Kodaira,H., Kondo,H.,				
Sugawara,M., Takabashi,M., Kanda,K., Yokoi,T., Furuya,T.,				
Kikkawa,E., Omura,Y., Abe,K., Kamihara,K., Katsuta,N., Sato,K.,				
Tanikawa,M., Yamazaki,M., Ninomiya,K., Ishibashi,T., Yamashita,H.,				
Murakawa,K., Fujimori,K., Tanai,H., Kimata,M., Watanabe,M.,				

QY 775 ileSerLysPheArgThrGlyLysIleAsnLeuLeuIleAlaThrThrValAlaGluGlu 794
|||||
Db 901 ATTAGTAAATTCGCACCTGGAAGATAAATCTGCTATCGTACCACAGTGGCAGAGAA 960

QY 795 GlyLeuAspIleLysGluCysAsnIleValIleArgTyrGlyLeuValThrAsnGluIle 814
|||||
Db 961 GGTCTGGATATTAAAGAATGTAAATGTTATCCGTTATGGTCTCGTCAACCAATGAAATA 1020

QY 815 AlaMetValGlnAlaArgGlyArgAlaArgAlaAspGluSerThrTyrValLeuValAla 834
|||||
Db 1021 GCCATGGTCCAGGCCCGTGGTCGAGCCAGAGCTGATGAGAGCACCTACGTCCTGGTTGCT 1080

QY 835 HisSerGlySerGlyValIleGluHisGluThrValAsnAspPheArgGluLysMetMet 854
|||||
Db 1081 CACAGTGGTTCCAGGAGTTATCGAACGTGAGACAGATTAAATGATTTCCGAGAGAAGATGAT 1140

QY 855 TyrLysAlaIleHisCysValGlnAsnMetLysProGluGluTyrAlaHisLysIleLeu 874
|||||
Db 1141 TATAAAGCTATACATTTGTTTCAAAATATGAACCCAGAGAGTATGCTCATAGATTTTG 1200

QY 875 GluLeuGlnMetGlnSerIleMetGluLysLysMetLysThrLysArgAsnIleAlaLys 894
|||||
Db 1201 GAATTACAGATGCAAGTATAATGGAAGAAATGAAGCAAGAGAAATATGCCAAG 1260

QY 895 HisTyrLysAsnAsnProSerLeuIleThrPheLeuCysLysAsnCysSerValLeuAla 914
|||||
Db 1261 CATTACAAGATAACCCATCACTAATACTTCTTTGCAAAACTGCAGTGTGTAGCC 1320

QY 915 CysSerGlyGluAspIleHisValIleGluLysMetHisHisValAsnMetThrProGlu 934
|||||
Db 1321 TGTTCCTGGGAAGATATCCATGTAAATGAGAAATGCATCAGTCAATATGACCCAGAA 1380

QY 935 PheLysGluLeuTyrIleValArgGluAsnLysAlaLeuGlnLysLysCysAlaAspTyr 954
|||||
Db 1381 TTCAAGGAACCTTTACATTTGTAAGAGAAACAAACACTGCAAAAGAAAGTGTGCCGACTAT 1440

QY 955 GlnIleAsnGlyGluIleIleCysLysCysGlyGlnAlaTrpGlyThrMetMetValHis 974
|||||
Db 1441 CAAATAAATGGTGAATCATCTGCAAAATGTGCCCAGGCTTGGGGAACAATGATGGTGCAC 1500

QY 975 LysGlyLeuAspLeuProCysLeuLysIleArgAsnPheValValPheLysAsnAsn 994
|||||
Db 1501 AAAGGCTTAGATTTGCCTTGCTCTCAAAATAAGGAATTTTGTAGTGGTTTCAAAAATAAT 1560

QY 995 SerThrLysLysGlnTyrLysLysTyrValGluLeuProIleThrPheProAsnLeuAsp 1014
|||||
Db 1561 TCAACAAAGAAACAATACAAAAAGTGGGTAGAATTACCTATCACATTTCCCAATCTTGAC 1620

QY 1015 TyrSerGluCysCysLeuPheSerAspGluAsp 1025
|||||
Db 1621 TATTCAAGATGCTGTTTATTATTAGTGATGAGGAT 1653

RESULT 12
BC025508
LOCUS BC025508 2468 bp mRNA linear ROD 20-JUL-2004
DEFINITION Mus musculus interferon induced with helicase C domain 1, mRNA
(cDNA clone IMAGE:5252257), complete cds.

ACCESSION BC025508
VERSION BC025508.1 GI:19343862
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 2468)
AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,

Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skaleka,U., Smailus,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 2468)
Strausberg,R.
Direct Submission
Submitted (06-MAR-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@nhgri.nih.gov
Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,
Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C.,
McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W.,
Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
Young,A., Zhang,L.-H. and Green,B.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 56 Row: c Column: 12
This clone has the following problem: The cds is short compared to
the longest cds in the locus.

FEATURES
source
1. .2468
/organism="Mus musculus"
/mol_type="mRNA"
/strain="mix FVB/N, C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:5252257"
/tissue_type="Mammary tumor. WAP-TGF alpha model. 7 months
old, gross tissue."
/clone_lib="NCI CGAP_Mam5"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1. .2468
/gene="Ifih1"
/note="synonyms: Helicard, Hlcd, MDA5, MDA-5"
/db_xref="LocusID:71586"
/db_xref="MGI:1918836"
444. .1847
/gene="Ifih1"
/codon_start=1
/product="Ifih1 protein"
/protein_id="AAH25508.1"
/db_xref="GI:19343863"
/db_xref="LocusID:71586"
/db_xref="MGI:1918836"
/translation="WASIQTYCQKSPMSDFGTQHYEQWAIQMEKKAADGNRKDRVCA

gene

CDS

EHLRKNEALQINDTIRMIDAHSHLETFYDTDEKEKFAVLNDSDESSNDQLK
GDVKSLKIDETDEFLMNLFPDNKMKLKLAEENPKYENEKLIKLRNTILSQFTRSEES
KRGIIFTKTRQSTYALSONMENAKFAEVGVKAHLIGAGHSSEVKPMTQTEQKEVIS
KFGTGEINLLIATTVAEEGLDIKECNIVIRYGLVTNFIAMVQARGRARADESTYVLVT
SSGSGVTEREIVNDFREKMYKAINRVQNMKPEEYAHKILELQVQSILEKMKVKRSI
AKQYNDNPSLITLLCKNCSMLVCSGENIHVIERMHVNMTPEFKGLYIVRENKALOKK
FADYQTNGEIICKGQAWGTMVHKGLDLPCLKIRNFVFNKNSPKQYKKWVELPI
RFPDLDYSEYCLYSDEN"

ORIGIN

Alignment Scores: 2.2e-173 Length: 2468
Pred. No.: 2690.50 Matches: 518
Score: 2690.50
Percent Similarity: 91.01% Conservative: 39
Best Local Similarity: 84.64% Mismatches: 54
Query Match: 50.66% Indels: 1
DB: 10 Gaps: 1

US-09-515-363C-2 (1-1025) x BC025508 (1-2468)

QY	414	AlaGlnIleLeuGluAsnSerLeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGln	433
Db	12	GCTCAAAATCCTTGAAAACTCCCTCTTAAATCTGGAGAGTGGAGACGATGACGGTGTGCAG	71
QY	434	LeuSerAspPheSerLeuIleIleIleAspGluCysHisHisThrAsnLysGluAlaVal	453
Db	72	CTGTCAGACTTCTCTCTCATTATCATTTGATGAGTGCCATCACACCAACAGGAGGCAGTC	131
QY	454	TyrAsnAsnIleMetArgHisTyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysLys	473
Db	132	TATAACAACATCATGACGATATTTTGAAGCAGAAGCTGAGAAACAATGACCTCAAGAAA	191
QY	474	GluAsnLysProValIleProLeuProGlnIleLeuGlyLeuThrAlaSerProGlyVal	493
Db	192	CAAAACAAACCCAGCCATTCCCCTGCCGCAGATACTAGGACTGACAGCTTCACCTGGTGT	251
QY	494	GlyGlyAlaThrLysGlnAlaLysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeu	513
Db	252	GGAGCAGCCAAAAGCAGTCTGAGGCTGAAAAACATATTTTAAATATATGTGCCAATCTT	311
QY	514	AspAlaPheThrIleLysThrValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGln	533
Db	312	GATGCCTTTACCATTAAAAACAGTGAAAGAGAAATCTTGGTCAACTCAAAACACCAATAAAG	371
QY	534	GluProCysLysLysPheAlaIleAlaAspAlaThrArgGluAspPropheLysGluLys	553
Db	372	GAACCATGCAAGAAATTGTGATTGCTGATGACACCAGAGAAATCCATTTAAAGAGAAA	431
QY	554	LeuLeuGluIleMetThrArgIleGlnThrTyrCysGlnMetSerProMetSerAspPhe	573
Db	432	CTTCTAGAAATTATGGCAAGCATTCAGACTTACTGCCAAAAAAGTCCAATGTCCAGATTT	491
QY	574	GlyThrGlnProTyrGluGlnTrpAlaIleGlnMetGluLysLysAlaAlaLysLysGly	593
Db	492	GGAACCCACATTATGAGCAGTGGGCCATTCAAAATGGAGAAAAAAGCTGCTAAAGACGGA	551
QY	594	AsnArgLysGluArgValCysAlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIle	613
Db	552	AATCGAAAGATCGGCTGTGCAGAACATTTGAGGAAGTACAAAGAGCCCTACAAATC	611
QY	614	AsnAspThrIleArgMetIleAspAlaTyrThrHisLeuGluThrPheTyrAsnGluGlu	633
Db	612	AACGACACGATCCGAATGATTGATGCACATAGCCACCTGGAGACATTCTACACTGATGAG	671
QY	634	LysAspLysLysPheAlaValIleGluAspAspSerAspGluGlyGlyAspAspGluTyr	653
Db	672	AAAGAAAAGAAGTTCGAGTCCTC---AATGACAGCGACGAGAGTGATGACGAGGCCAGC	728
QY	654	CysAspGlyAspGluAspGluAspAspLeuLysLysProLeuLysLeuAspGluThrAsp	673
Db	729	AGTTGCAATGACCAACTTAAGGGCGATGTAAAGAAAAATCTTTGAAACTGGACGAAACGGAT	788
QY	674	ArgPheLeuMetThrLeuPhePheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsn	693

Db	789	GAATTTCTCATGAATTTGTTCTTTGATAACAAGAAAATGTTGAAAAAACTAGCTGAAAC	848
QY	694	ProGluTyrGluAsnGluLysLeuThrLysLeuArgAsnThrIleMetGluGlnTyrThr	713
Db	849	CCAAAATACGAGAATGAAAAACTCATTAATTAAGAAACACGATACTGGAACAATTCACA	908
QY	714	ArgThrGluGluSerAlaArgGlyIleIlePheThrLysThrArgGlnSerAlaTyrAla	733
Db	909	AGGTCTGAGGAGTCCTCCCGAGGAATTATTTTACAAAAACACGACAGACCTACGCA	968
QY	734	LeuSerGlnTrpIleThrGluAsnGluLysPheAlaGluValGlyValLysAlaHisHis	753
Db	969	CTTTCCCACTGGATCATGGAATAATGCAAAAGTTTTCGGAGTTGGAGTCAAAGCGCATCAC	1028
QY	754	LeuIleGlyAlaGlyHisSerSerGluPheLysProMetThrGlnAsnGluGlnLysGlu	773
Db	1029	CTGATTGGCGGGGCACAGCAGTGAAGTCAAGCCCCATGACTCAGACTGAACAAAAAGAA	1088
QY	774	ValIleSerLysPheArgThrGlyLysIleAsnLeuLeuIleAlaThrThrValAlaGlu	793
Db	1089	GTCATTAGTAAATTTTCGCACTGGCGAAATAAATCTGCTTATCGTACGACGGTGGCAGAG	1148
QY	794	GluGlyLeuAspIleLysGluCysAsnIleValIleArgTyrGlyLeuValThrAsnGlu	813
Db	1149	GAAGGCCTGGATATCAAAGAGTGCAATATTGTTATTTCGTTATGGCCTTGTACGAACGAG	1208
QY	814	IleAlaMetValGlnAlaArgGlyArgAlaArgAlaAspGluSerThrTyrValLeuVal	833
Db	1209	ATAGCCATGGTCCAGGCCCGGGTCGAGCCAGAGCTGATGAAAGCACGTATGTCTCTGGTC	1268
QY	834	AlaHisSerGlySerGlyValIleGluHisGluThrValAsnAspPheArgGluLysMet	853
Db	1269	ACCAGCAGTGGCTCAGGAGTTACCGAACGGGAGATTGTTAATGATTTCCGAGAGAAGATG	1328
QY	854	MetTyrLysAlaIleHisCysValGlnAsnMetLysProGluGluTyrAlaHisLysIle	873
Db	1329	ATGTATAAAGCTATTAAACCGTGTTCAAACACATGAAACCAGAGGAGTATGCACATAAGAT	1388
QY	874	LeuGluLeuGlnMetGlnSerIleMetGluLysLysMetLysThrLysArgAsnIleAla	893
Db	1389	TTGGAATTGCAGGTGCAAAAGTATCTCTGGAAAAAGAAAATGAAAGTCAAAGAACGATTGCA	1448
QY	894	LysHisTyrLysAsnAsnProSerIleuIleThrPheLeuCysLysAsnCysSerValLeu	913
Db	1449	AAGCAATACAACGACAATCCATCGTTAATAAACACTTCTCTGCAAAAAATTTGTAGCATGCTG	1508
QY	914	AlaCysSerGlyGluAspIleHisValIleGluLysMetHisHisValAsnMetThrPro	933
Db	1509	GTCTGCTCGGAGAAAAACATCCATGTTCATTGAGAAGATGCATCATGTCAATATGACACCA	1568
QY	934	GluPheLysGluLeuTyrIleValArgGluAsnLysAlaLeuGlnLysCysAlaAsp	953
Db	1569	GAATTCAAGGGACTCTACATTGTGAAGAGAAAAACAAAGCACTGCAAAAGAAAATTTGCTGAT	1628
QY	954	TyrGlnIleAsnGlyGluIleIleCysLysCysGlyGlnAlaTrpGlyThrMetMetVal	973
Db	1629	TATCAGACCAATGGAGAGATTATCTGCAAGTGTGGCCAGGCTTGGGGAACAATGATGGTG	1688
QY	974	HisLysGlyLeuAspLeuProCysLeuLysIleArgAsnPheValValPheLysAsn	993
Db	1689	CACAAAGGTTTAGATTTGCTTGTAAAATAAGGAATTTTGTAGTCAATTTTCAAAAAAT	1748
QY	994	AsnSerThrLysLysGlnTyrLysLysTrpValGluLeuProIleThrPheProAsnLeu	1013
Db	1749	AACTCACCGAAGAAACAGTACAGAAGTGGGTGGAATTTGCCTATCAGATTTTCCTGATCTT	1808
QY	1014	AspTyrSerGluCysCysLeuPheSerAspGluAsp	1025
Db	1809	GACTACTCAGAATACTGCTTGTATAGTATGAAGAT	1844

RESULT 13
AX300841

LOCUS AX300841 1443 bp DNA linear PAT 30-NOV-2001
DEFINITION Sequence 10 from Patent WO0185955.
ACCESSION AX300841
VERSION AX300841.1 GI:17382119
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Bahr, G., Cocude, C. and Capron, A.
TITLE Rh116 polypeptides and its fragments and polynucleotides encoding
said polypeptides and therapeutic uses
JOURNAL Patent: WO 0185955-A 10 15-NOV-2001;
Istac (FR) ; INSTITUT PASTEUR DE LILLE (FR)
FEATURES
source Location/Qualifiers
1. .1443
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 8.88e-157 Length: 1443
Score: 2442.00 Matches: 475
Percent Similarity: 99.38% Conservative: 2
Best Local Similarity: 98.96% Mismatches: 3
Query Match: 45.98% Indels: 0
DB: 6 Gaps: 0
US-09-515-363C-2 (1-1025) x AX300841 (1-1443)
Qy 174 LysGluAsnTrpPheSerAlaPheLeuAsnValLeuArgGlnThrGlyAsnAsnGluLeu 193
Db 2 AAAGAAACTGGTTCCTGCAATTCCTGAATGTTCTTCGTCACAGGAAACAATGAACCT 61
Qy 194 ValGlnGluLeuThrGlySerAspCysSerGluSerAsnAlaGluIleGluAsnLeuSer 213
Db 62 GTCCAAGAGTTAACAGGCTCTGATTGCTCAGAAAGCAATGCAGAGATTGAGAAATTATCA 121
Qy 214 GlnValAspGlyProGlnValGluGlnLeuLeuSerThrThrValGlnProAsnLeu 233
Db 122 CAAGTTGATGGTCCCTCAAGTGGAGAGCAACTTCTTCAACCCACAGTTCAGCCAAATCTG 181
Qy 234 GluLysGluValTrpGlyMetGluAsnAsnSerSerGluSerSerPheAlaAspSerSer 253
Db 182 GAGAAGGAGGTCTGGGGCATGGAGAAATAACTCATCAGAATCATCTTTTCAGATTCCTCT 241
Qy 254 ValValSerGluSerAspThrSerLeuAlaGluGlySerValSerCysLeuAspGluSer 273
Db 242 GTAGTTTCAGAAATCAGACACAAGTTTGGCAGAAGGAAGTGTGAGCTGCTTAGATGAAAGT 301
Qy 274 LeuGlyHisAsnSerAsnMetGlySerAspSerGlyThrMetGlySerAspSerAspGlu 293
Db 302 CTTGGACATAACAGCAACATGGGCAGTGATTTCAGGCACCATGGGAAGTGATTTCAGATGAA 361
Qy 294 GluAsnValAlaAlaArgAlaSerProGluProGluLeuGlnLeuArgProTyrGlnMet 313
Db 362 GAGAAATGGCAGCAAGAGCATCCCGGAGCCAGAACTCCAGCTCAGGCCTTACCAAATG 421
Qy 314 GluValAlaGlnProAlaLeuGluGlyLysAsnIleIleIleCysLeuProThrGlySer 333
Db 422 GAAGTTGCCAGCCAGCCTTGGAGGGGAAGAATATCATCATCTGCCTCCCTACAGGGAGT 481
Qy 334 GlyLysThrArgValAlaValTyrIleAlaLysAspHisLeuAspLysLysLysAla 353
Db 482 GGAAAAACCAGAGTGGCTGTTTACATTGCCAAGGATCACTTAGACAAGAAAGCA 541
Qy 354 SerGluProGlyLysValIleValLeuValAsnLysValLeuLeuValGlnLeuPhe 373
Db 542 TCTGAGCCTGGAAAAGTTATAGTCTTGTCAATAAGGTACTGCTAGTTGAACAGCTCTTC 601
Qy 374 ArgLysGluPheGlnProPheLeuLysLysTrpTyrArgValIleGlyLeuSerGlyAsp 393

Db 602 CGCAAGGAGTTCCAACCAATTTTGAAGAAATGGTATCGTGTATTGGATTAAAGTGGTGAT 661
Qy 394 ThrGlnLeuLysIleSerPheProGluValValLysSerCysAspIleIleSerThr 413
Db 662 ACCCAACTGAAATATCATTTCCAGAAAGTTGTCAAGTCTCTGTGATATTATTCAGTACA 721
Qy 414 AlaGlnIleLeuGluAsnSerLeuAsnLeuGluAsnGlyGluAspAlaGlyValGln 433
Db 722 GCTCAATCCTTGAAAACTCCCTCTTAAACTTGGAAAAATGGAGAAGATGCTGGTGTCAA 781
Qy 434 LeuSerAspPheSerLeuIleIleIleAspGluCysHisHisThrAsnLysGluAlaVal 453
Db 782 TTGTCAGACTTTTCCCTTCATTATTCATTGATGAATGTCATCACCAACAAGAACGAGTG 841
Qy 454 TyrAsnAsnIleMetArgHisTyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysLys 473
Db 842 TATAATAACATCATGAGGCATTATTGATGCAGAAAGTTGAAAAACAATAGACTCAAGAA 901
Qy 474 GluAsnLysProValIleProLeuProGlnIleLeuGlyLeuThrAlaSerProGlyVal 493
Db 902 GAAAAACAAACCAAGTGATTCCCTTCNTCAGATACTGGGACTAACAGCTTCACCTGGTGT 961
Qy 494 GlyGlyAlaThrLysGlnAlaLysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeu 513
Db 962 GGAGGGGCCACGAAAGCAAGCCAAAGCTGAAGAACACACATTTTAAAACTATGTGCCAATCTT 1021
Qy 514 AspAlaPheThrIleLysThrValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGln 533
Db 1022 GATGCATTTACTATTAAACTGTTTAAAGAAAAACCTTGATCAACTGAAAAACCAAAATACAG 1081
Qy 534 GluProCysLysLysPheAlaIleAlaAspAlaThrArgGluAspPropheLysGluLys 553
Db 1082 GAGCCATGCAAGAGTTTGCCATTGCAGATGCAACCAGAGAAGATCCATTTTAAAGAGAAA 1141
Qy 554 LeuLeuGluIleMetThrArgIleGlnThrTyrCysGlnMetSerProMetSerAspPhe 573
Db 1142 CTTCTAGAAATATGACAAGGATTCAAACTTATTGTCAAATGAGTCCAATGTCCAGATTTT 1201
Qy 574 GlyThrGlnProTyrGluGlnTrpAlaIleGlnMetGluLysLysAlaAlaLysLysGly 593
Db 1202 GGAACCTCAACCTATGAACAATGGGCCATTCAAATGGAAAAAAGCTGCAAAAGAGGA 1261
Qy 594 AsnArgLysGluArgValCysAlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIle 613
Db 1262 AATCGAAAGAAAGTGTGTCAGAACATTGAGGAAGTACATAAGGCCCTACAAATT 1321
Qy 614 AsnAspThrIleArgMetIleAspAlaTyrThrHisLeuGluThrPheTyrAsnGluGlu 633
Db 1322 AATGACACAATTCGAATCATAGATGCGTATACTCATCTTGAAACTTCTATATGAAGAG 1381
Qy 634 LysAspLysLysPheAlaValIleGluAspAspSerAspGluGlyAspAspGluTyr 653
Db 1382 AAAGATAAGAAAGTTTGCAGTCATAGAAGATGATAGTAGGAGGTGGTGATGATGAGTAT 1441
RESULT 14
BC004031 2050 bp mRNA linear ROD 20-JUL-2004
LOCUS Mus musculus interferon induced with helicase C domain 1, mRNA
DEFINITION (cDNA clone IMAGE:3495361), partial cds.
ACCESSION BC004031
VERSION BC004031.1 GI:13278455
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 2050)
REFERENCE
AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Schuler, G.D.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, N.K.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,

Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettaman, M., Madan, A., Rodrigues, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smailus, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 2050)
Strausberg, R.
Direct Submission
Submitted (28-FEB-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgaps-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

TITLE

JOURNAL
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

REMARK
COMMENT

FEATURES
source

gene

CDS

ORIGIN
Alignment Scores:
Pred. No.: 5.1e-145 Length: 2050
Score: 2271.50 Matches: 438
Percent Similarity: 90.87% Conservative: 30
Best Local Similarity: 85.05% Mismatches: 46
Query Match: 42.77% Indels: 1
DB: 10 Gaps: 1
US-09-515-363C-2 (1-1025) x BC004031 (1-2050)
QY 511 AlaAsnLeuAspAlaPheThrIleLysThrVallysGluAsnLeuAspGlnLeuLysAsn 530
Db 1 GCCAATCTTGATGCCTTTACCATTAAACAGTGAAAGAGAAATCTTGGTCAACTCAACAC 60
QY 531 GlnIleGlnGluProCysLysLysPheAlaIleAlaAspAlaThrArgGluAspProphe 550
Db 61 CAAATAAGGAACCATGCAAGAAATTTGTGATTGCTGTATACCATCCAGAGAAATCCATT 120
QY 551 LysGluLysLeuLeuGluIleMetThrArgIleGlnThrTyrCysGlnMetSerProMet 570
Db 121 AAAGAGAAACTACTAGAAATTTATGGCAAGCATTCAGACTTACTGCCAAAAAAGTCCAATG 180
QY 571 SerAspPheGlyThrGlnProTyrGluGlnTrpAlaIleGlnMetGluLysLysAlaAla 590
Db 181 TCAGATTTTGGAAACCAACATTATAGCAGTGGCCATTCAAATGGAGAAAAAGCTGCT 240
QY 591 LysLysGlyAsnArgLysGluArgValCysAlaGluHisLeuArgLysTyrAsnGluAla 610
Db 241 AAAGACGGAATCGCAAAGATCGCGTCTGTGCAGAACATTTGAGGAAGTACACGAAGCC 300
QY 611 LeuGlnIleAsnAspThrIleArgMetIleAspAlaTyrThrHisLeuGluThrPheTyr 630
Db 301 CTACAAATCAACGACACGATCCGAATGATTGATGCGTATAGCCACCTGGAGACATTCTAC 360
QY 631 AsnGluGluLysAspLysLysPheAlaValIleGluAspAspSerAspGluGlyGlyAsp 650
Db 361 ACTGATGAGAAAGAAAGAAAGTTCCGAGTCCTC--AACGACAGCGACGAGAGTGATGAC 417
QY 651 AspGluTyrCysAspGlyAspGluAspGluAspAspLeuLysLysProLeuLysLeuAsp 670
Db 418 GAGGCCAGCAGTTGCAATGACCAACTTAAGGGCGATGTAAGAGAAATCTTTGAAACTGGAC 477
QY 671 GluThrAspArgPheLeuMetThrLeuPhePheGluAsnAsnLysMetLeuLysArgLeu 690
Db 478 GAAACGGATGAATTTCTCATGAATTTGTTCTTTGATAACAAGAAATGTTGAAAAAACTA 537
QY 691 AlaGluAsnProGluTyrGluAsnGluLysLeuThrLysLeuArgAsnThrIleMetGlu 710
Db 538 GCTGAAACCCAAATACGAGAAATGAAAAACTCATTAATTAAGAAACACGATACTGAA 597
QY 711 GlnTyrThrArgThrGluGluSerAlaArgGlyIleIlePheThrLysThrArgGlnSer 730
Db 598 CAATTCAAGGTCTGAGGAGTCTCCCGAGGAATTTATTTTCACAAAAACACGACAGAC 657
QY 731 AlaTyrAlaLeuSerGlnTrpIleThrGluAsnGluLysPheAlaGluValGlyValLys 750
Db 658 ACCTACGCACTTTCAGTGGATCATGGAAAAATGCAAGTTTGGGAAGTTGGAGTCAAA 717
QY 751 AlaHisHisLeuIleGlyAlaGlyHisSerSerGluPheLysProMetThrGlnAsnGlu 770
Db 718 GCGCATCACCTGATTGGCGCGGCGCACAGCAGTGAAGTCAAGCCCATGACTCAGACTGAA 777
QY 771 GlnLysGluValIleSerLysPheArgThrGlyLysIleAsnLeuLeuIleAlaThrThr 790
Db 778 CAAAAAGAGTCATTAGTAAATTTGCACTGGCGGAAATAAATCTGCTTATCGCTACGACG 837
QY 791 ValAlaGluGluGlyLeuAspIleLysGluCysAsnIleValIleArgTyrGlyLeuVal 810

LEIMASITQYCKSPMSDFGTHYEQWAIQMEKKAADGNRKDRVCAEHLRKYNEALQ
INDTRIMDAYSHLETFTDEKEKKFAVLNDSDESDEASSCNDQLKGDVKSLKLDE
TDEFMLNLFDDNKKMLKLAENPKYENEKLIKLRNILEQFTRSESSRGIIFTKTRQ
STYALSQWIMENAKFAEVGVKAHLHIGAGHSSEVPMTQTEQKEVISKFRFTGEINLLI
ATTVAEGLDIKECNIVIRYGLVTNEIAMVQARGARADESTYVLVTSSGSGVTEREI
VNDFREKMYKAINRVQNMKPBEYAHKILELQVQSILEKKMKVKRSIAQYNDNPSLI

Db 838 GTGGCAGAGGAGGCTGGATATCAAAAGAGTGCAATATTGTTATTTCGTTATGGCCTTGTC 897

QY 811 ThrAsnGluIleAlaMetValGlnAlaArgGlyArgAlaArgAlaAspGluSerThrTyr 830

Db 898 ACGAATGAGATAGCCATGGTCCAGGCCCGGGTCGAGCCAGAGCTGATGAAGCACGTAC 957

QY 831 ValLeuValAlaHisSerGlySerGlyValIleGluHisGluThrValAsnAspPheArg 850

Db 958 GTCCTGGTCACCGACAGTGGCTCAGGAGTTACCGAACCGGAGATTGTTAATGATTCCGA 1017

QY 851 GluLysMetMetTyrLysAlaIleHisCysValGlnAsnMetLysProGluGluTyrAla 870

Db 1018 GAGAAGATGATGTATAAAGCTATTAAACCGTGTTCAAAACATGAACACGAGGAGTATGCA 1077

QY 871 HisLysIleLeuGluLeuGlnMetGlnSerIleMetGluLysLysMetLysThrLysArg 890

Db 1078 CATAAGATTTTGGAAATTGCAGGTGCAAAAGTATCCTGGAAAAAGAAAATGAAAGTCAAAAAGA 1137

QY 891 AsnIleAlaLysHisTyrLysAsnAsnProSerLeuIleThrPheLeuCysLysAsnCys 910

Db 1138 AGCATTGCAAAAGCAATACAACGACATCCATCGTTAATAACACTTCTCTGCAAAAATTGT 1197

QY 911 SerValLeuAlaCysSerGlyGluAspIleHisValIleGluLysMetHisHisValAsn 930

Db 1198 AGCATGCTGGTCTGCTCGGAGAGAAACATCCATGTCATTGAGAAGATGTCATCATGTCAAT 1257

QY 931 MetThrProGluPheLysGluLeuTyrIleValArgGluAsnLysAlaLeuGlnLysLys 950

Db 1258 ATGACACCCAGAATTCAAGGACTCTACATTGTTAAGAGAAAAACAAAGCACCTGCAAAAAGAAA 1317

QY 951 CysAlaAspTyrGlnIleAsnGlyGluIleIleCysLysCysGlyGlnAlaTrpGlyThr 970

Db 1318 TTTGCTGATTATCAGACCAATGGAGAGATTATCTGCAAGTGTGGCCAGGCTTGGGGAACA 1377

QY 971 MetMetValHisLysGlyLeuAspLeuProCysLeuLysIleArgAsnPheValValVal 990

Db 1378 ATGATGGTGCAAAAGGTTTAGATTTCCTTGTCTTAAAAAAGGAATTTTGTAGTCAAT 1437

QY 991 PheLysAsnAsnSerThrLysLysGlnTyrLysLysTyrValGluLeuProIleThrPhe 1010

Db 1438 TTCAAAAATACTCACCGAAGAACAGTACAAAGAGTGGGTGGAATTGCTTATCAGATT 1497

QY 1011 ProAsnLeuAspTyrSerGluCysCysLeuPheSerAspGluAsp 1025

Db 1498 CCTGATCTTGACTACTCAGAAATACTGCTTGATATAGTGAAGAT 1542

RESULT 15

AR379503

LOCUS AR379503 1392 bp DNA linear PAT 18-DEC-2003

DEFINITION Sequence 48 from patent US 6607879.

ACCESSION AR379503

VERSION AR379503.1 GI:40087137

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1392)

AUTHORS Cocks,B.G., Stuart,S.G. and Seilhamer,J.J.

TITLE Compositions for the detection of blood cell and immunological response gene expression

JOURNAL Patent: US 6607879-A 48 19-AUG-2003;

FEATURES

source 1. .1392

/organism="unknown"

/mol_type="genomic DNA"

ORIGIN

Alignment Scores:

Pred. No.: 1.15e-131 Length: 1392

Score: 2072.00 Matches: 400

Percent Similarity: 94.79% Conservative: 0

Best Local Similarity: 94.79% Mismatches: 22

Query Match: 39.01% Indels: 0

DB: 6 Gaps: 0

US-09-515-363C-2 (1-1025) x AR379503 (1-1392)

QY 604 LeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMetIleAspAlaTyr 623

Db 3 TTGAGGAAGTACATGAGGCCCTACAAATTAATGACACAATTCGAATGATAGATGCGTAT 62

QY 624 ThrHisLeuGluThrPheTyrAsnGluGluLysAspLysLysPheAlaValIleGluAsp 643

Db 63 ACTCATCTTGAAACCTTCTATATAATGAAGAGAGAAAGATAAGAAAGTTTGCACTCATANNNNN 122

QY 644 AspSerAspGluGlyAspAspGluTyrCysAspGlyAspGluAspGluAspLeu 663

Db 123 NNTTA 182

QY 664 LysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeuPhePheGluAsn 683

Db 183 AAGAAAACCTTTGAAACTGGATGAAACAGATAGATTCTCATGACTTTATTTTGGAAAAC 242

QY 684 AsnLysMetLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGluLysLeuThrLys 703

Db 243 AATAAAATGTTGAAAAGCTGGCTGAAAACCCAGAAATATGAAAATGAAAAAGCTGACCAA 302

QY 704 LeuArgAsnThrIleMetGluGlnTyrThrArgThrGluGluSerAlaArgGlyIleIle 723

Db 303 TTAAGAAATACCATTAATGGAGCAATATACTAGGACTGAGGAATCAGACGAGGAATAATC 362

QY 724 PheThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTrpIleThrGluAsnGluLys 743

Db 363 TTTCAAAAACACGACAGAGTGCAATATGCGCTTTCCAGTGGATTACTGAAAATGAAAAA 422

QY 744 PheAlaGluValGlyValLysAlaHisHisLeuIleGlyAlaGlyHisSerSerGluPhe 763

Db 423 TTTGCTGAAGTAGGAGTCAAAGCCCAACCATCTGATTGGAGTGGACACAGCAGTGAGTTC 482

QY 764 LysProMetThrGlnAsnGluGlnLysGluValIleSerLysPheArgThrGlyLysIle 783

Db 483 AAACCCATGACACAGAAATGAACAAAAAGAAAGTCATTAGTAAATTTCCGACTGGAANAATA 542

QY 784 AsnLeuLeuIleAlaThrThrValAlaGluGluGlyLeuAspIleLysGluCysAsnIle 803

Db 543 AATCTGCTTATCGCTACCACAGTGGCAGAGAGGTCCTGGATATTAAAGAAATGTAACATT 602

QY 804 ValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAlaArgGlyArgAla 823

Db 603 GTTATCCGTTATGGTCTCGTCACCAATGAATAGCCATGGTCCAGGCCCGCTGGTCGAGCC 662

QY 824 ArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGlyValIleGluHis 843

Db 663 AGAGCTGATGAGAGCACCTACGTCCTGGTTGCTCACAGTGGTTCAGGAGTTATCGAACGT 722

QY 844 GluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHisCysValGlnAsn 863

Db 723 GAGACAGTTAATGATTTCCGAGAGAGAGATGATGTATAAAGCTATACATTGTGTTCAAAAT 782

QY 864 MetLysProGluGluTyrAlaHisLysIleLeuGluLeuGlnMetGlnSerIleMetGlu 883

Db 783 ATGAAACCAGAGGAGTATGCTCATAAGATTTTGGAAATTACAGATGCAAAAGTATAATGGAA 842

QY 884 LysLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAsnProSerLeuIle 903

Db 843 AAGAAAATGAAACCAAGAGAAATATTGCCAAGCATTACAAGAATAACCCATCACTAATA 902

QY 904 ThrPheLeuCysLysAsnCysSerValLeuAlaCysSerGlyGluAspIleHisValIle 923

Db 903 ACTTTCCTTTGCAAAAACCTGCAGTGTGCTAGCCCTGTTCTGGGGAAGATATCCATGTAATT 962

QY 924 GluLysMetHisHisValAsnMetThrProGluPheLysGluLeuTyrIleValArgGlu 943

Db 963 GAGAAAATGCATCACGTCATATATGACCCCAAGAAATTCAGGAACCTTTACATTGTAAGAGAA 1022

QY 944 AsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluIleIleCysLys 963

Db	1023	AACAAAGCACTGCAAAAGAAGTGTGCCGACTATCAAAATAAATGGTGAATCATCTGCAAA	1082
QY	964	CysGlyGlnAlaTrpGlyThrMetMetValHisLysGlyLeuAspLeuProCysLeuLys	983
Db	1083	TGTGGCCAGGCTTGGGGAAACAATGATGGTGCAACAAGGCTTAGATTTGCCTTGTCTCAAA	1142
QY	984	IleArgAsnPheValValPheLysAsnAsnSerThrLysLysGlnTyrLysLysTrp	1003
Db	1143	ATAAGGAATTTTGTAGTGGTTTTCAAAAATAATTCAACAAGAAACAATACAAAAAGTGG	1202
QY	1004	ValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCysCysLeuPheSerAsp	1023
Db	1203	GTAGAAATTACCTATCACATTTCCCAATCTTGACTATTGAGAATGCTGTTTATTAGTGAT	1262
QY	1024	GluAsp 1025	
Db	1263	GAGGAT 1268	

Search completed: January 28, 2005, 00:36:14
Job time : 9477 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 27, 2005, 17:52:27 ; Search time 1011 Seconds
(without alignments)
5322.113 Million cell updates/sec

Title: US-09-515-363C-2
Perfect score: 5311
Sequence: 1 MSGXSTDENFRYLISCFRA.....LPITFPNLDYSECCLFSD 1025

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09515363/runat_26012005_152042_6131/app_query.fasta_1.1223
-DB=N_Geneseq_23Sep04 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09515363 @CGN_1_1_708 @runat_26012005_152042_6131 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :		N_Geneseq_23Sep04:*	
1:	geneseqn1980s:*		
2:	geneseqn1990s:*		
3:	geneseqn2000s:*		
4:	geneseqn2001as:*		
5:	geneseqn2001bs:*		
6:	geneseqn2002as:*		
7:	geneseqn2002bs:*		
8:	geneseqn2003as:*		
9:	geneseqn2003bs:*		
10:	geneseqn2003cs:*		
11:	geneseqn2003ds:*		
12:	geneseqn2004s:*		

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	5311	100.0	3365	4	AAD17203
2	5311	100.0	3380	12	ADJ74904
3	5285	99.5	3372	6	ABA04908
4	5285	99.5	3446	10	ADC30823
5	5283	99.5	3379	12	ADN04879
6	5276.5	99.4	3131	4	AAD11170

ALIGNMENTS

RESULT 1		AAD17203	
ID	AAD17203	standard; cDNA; 3365 BP.	
XX			
AC	AAD17203;		
XX			
DT	29-NOV-2001	(first entry)	
XX			
DE	Human melanoma differentiation associated (Mda)-5 cDNA.		
XX			
KW	Human; melanoma differentiation associated gene; Mda-5; interferon; IFN;		
KW	RNA helicase motif; caspase recruitment domain; CARD; therapy; melanoma;		
KW	neuroblastoma; astrocytoma; glioblastoma; multiforme; cancer; cervical;		
KW	breast; colon; prostate; osteosarcoma; chondrosarcoma; systemic toxicity;		
KW	central nervous system; cytosolic; apoptosis; ss.		
XX			
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
CDS		169..3246	
FT		/*tag= a	
FT		/product= "Human melanoma differentiation associated	
FT		(Mda)-5 protein"	
FT	misc_feature	3225..3228	
FT		/*tag= b	
FT		/note= "ATTTA motif"	
FT	misc_feature	3284..3287	
FT		/*tag= c	
FT		/note= "ATTTA motif"	
FT	polyA_signal	3343..3348	

QY 481 LeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyAlaThrLysGlnAla 500
Db 1609 CTTCTCAGACTGGGACTAACAGCTTCACCTGGTGTGGAGGGGCCACGAAGCAAGCC 1668
QY 501 LysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThrIleLysThr 520
Db 1669 AAAGCTGAAGAACACATTTTAAACTATGTGCCAATCTTGATGCATTTACTATTAAACT 1728
QY 521 VallysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluProCysLysLysPheAla 540
Db 1729 GTTAAAGAAAACCTTGATCAACTGAANAACCAATATACAGAGCCATGCAAGAGTTGCC 1788
QY 541 IleAlaAspAlaThrArgGluAspProPheLysGluLysLeuLeuGluIleMetThrArg 560
Db 1789 ATTGCAGATCAACACAGAGAGATCCATTTAAAGAGAACTTCTAGAAATAATGACAAGG 1848
QY 561 IleGlnThrTyrCysGlnMetSerProMetSerAspPheGlyThrGlnProTyrGluGln 580
Db 1849 ATTCAAACTTATGTCAAATGAGTCCAATGTCCAGATTTTGGAACTCAACCCCTATGAACAA 1908
QY 581 TrpAlaIleGlnMetGluLysLysAlaAlaLysLysGlyValAsnArgLysGluArgValCys 600
Db 1909 TGGGCCATTCAAATGGAAAAAAGCTGCAAAAAGGAATCGCAAGAACGTGTTGT 1968
QY 601 AlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMetIle 620
Db 1969 GCAGAACATTTGAGGAAGTACAAATGAGGCCCTACAAATTAATGACACAAATTCGAATGATA 2028
QY 621 AspAlaTyrThrHisLeuGluThrPheTyrAsnGluGluLysAspLysLysPheAlaVal 640
Db 2029 GATGCGTATACTCATCTTGAAACTTTCTATAATGAAGAGAAAGATAAGAAGTTTGCAGTC 2088
QY 641 IleGluAspAspSerAspGluGlyGlyAspAspGluTyrCysAspGlyAspGluAspGlu 660
Db 2089 ATAGAAGATGATAGTATGATGAGGGTGGTATGATGAGTATTGTGATGGTATGAAGATGAG 2148
QY 661 AspAspLeuLysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeuPhe 680
Db 2149 GATGATTTAAAGAAACCTTTGAAACTGGATGAACAGATAGATTTCTCATGACTTTATT 2208
QY 681 PheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGluLys 700
Db 2209 TTTGAAAACATAAAATGTTGAAAAGGCTGGCTGAAACCAGAAATATGAAAATGAAAAG 2268
QY 701 LeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArgThrGluGluSerAlaArg 720
Db 2269 CTGACCAAATTAAGAAATACCATAATGGAGCAATATACTAGGACTGAGGAATCAGCACGA 2328
QY 721 GlyIleIlePheThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTrpIleThrGlu 740
Db 2329 GGAATAATCTTTACAAAAACACGACAGAGTGCATATGCGCTTCCAGTGGATTACTGAA 2388
QY 741 AsnGluLysPheAlaGluValGlyValLysAlaHisHisLeuIleGlyAlaGlyHisSer 760
Db 2389 AATGAAAAATTTGCTGAAGTAGGAGTCAAAAGCCCACCATCTGATTGGAGCTGGACACAGC 2448
QY 761 SerGluPheLysProMetThrGlnAsnGluGlnLysGluValIleSerLysPheArgThr 780
Db 2449 AGTGAGTTCAACCCCATGACACAGATGAACAAAAGAGTCATTAGTAAATTTCCGACT 2508
QY 781 GlyLysIleAsnLeuLeuIleAlaThrThrValAlaGluGluGlyLeuAspIleLysGlu 800
Db 2509 GGAAAAATCAATCTGCTTATCGCTACCACAGTGGCAGAGAGGCTCTGGATATTAAAGAA 2568
QY 801 CysAsnIleValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAlaArg 820
Db 2569 TGTAACATTGTTATCCGTTATGGTCTCGTCACCAATGAATAGCCATGGTCCAGGCCCGT 2628
QY 821 GlyArgAlaArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGlyVal 840
Db 2629 GGTGAGCCAGAGCTGATGAGAGCACTAGTCTCGTCTGTTGCTCACAGTGGTTCAGGAGTT 2688

QY 841 IleGluHisGluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHisCys 860
Db 2689 ATCGAACATGAGACAGTTAATGATTTCCGAGAGAAGATGATGTATAAAGCTATACATTGT 2748
QY 861 ValGlnAsnMetLysProGluTyrAlaHisLysIleLeuGluLeuGlnMetGlnSer 880
Db 2749 GTTCAAAATATGAACCAGAGAGTATGCTCATAAGATTTTGGAATTACAGATGCAAAAGT 2808
QY 881 IleMetGluLysLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAsnPro 900
Db 2809 ATAATGGAAAAAGAAATGAAAACCAAGAGAAATATTGCCAAGCATTTACAAGAATAACCCA 2868
QY 901 SerLeuIleThrPheLeuCysLysAsnCysSerValLeuAlaCysSerGlyGluAspIle 920
Db 2869 TCACATAATACTTTCCTTTGGCAAAACTGCAGTGTGTAGCCTGTTCTGGGGAAGATATC 2928
QY 921 HisValIleGluLysMetHisHisValAsnMetThrProGluPheLysGluLeuTyrIle 940
Db 2929 CATGTAAATTGAGAAAATGCATCAGTCAATATGACCCAGAAATTCAGGAACCTTTACATT 2988
QY 941 ValArgGluAsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluIle 960
Db 2989 GTAAGAGAAAAACAAGCACTGCAAAAAGAGTGTGCCGACTATCAAATNAATGGTGAATC 3048
QY 961 IleCysLysCysGlyGlnAlaTrpGlyThrMetMetValHisLysGlyLeuAspLeuPro 980
Db 3049 ATCTGCAAATGTGGCCAGGCTTGGGGAACAATGATGGTGCACAAAGGCTTAGATTGCTT 3108
QY 981 CysLeuLysIleArgAsnPheValValPheLysAsnAsnSerThrLysLysGlnTyr 1000
Db 3109 TGTCTCAAAATAAGGAATTTTGTAGTGGTTTTTCAAAAATAATTCAACAAAGAAACAATAC 3168
QY 1001 LysLysTrpValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCysLeu 1020
Db 3169 AAAAGTGGGTAGAAATTACCTATCACATTTCCCAATCTTGACTATTTCAGAAATGCTGTTA 3228
QY 1021 PheSerAspGluAsp 1025
Db 3229 TTTAGTGATGAGGAT 3243
RESULT 2
ADJ74904
ID ADJ74904 standard; DNA; 3380 BP.
XX AC ADJ74904;
XX DT 20-MAY-2004 (first entry)
XX DE Marker gene SEQ ID NO:156.
XX KW bronchial asthma; chronic obstructive pulmonary disease;
KW KW respiratory epithelial cell; interleukin-13; respiratory; antiasthmatic;
KW KW gene therapy; marker gene; gene; ds.
XX OS Homo sapiens.
XX PN EP1394274-A2.
XX PD 03-MAR-2004.
XX PF 04-AUG-2003; 2003EP-00254857.
XX PR 06-AUG-2002; 2002JP-00229312.
XX PR 20-MAR-2003; 2003JP-00077212.
XX PA (GENO-) GENOX RES INC.
XX PI Ohtani N, Sugita Y, Yamaya M, Kubo H, Nagai H, Izuhara K;
XX DR WPI; 2004-193155/19.
XX PT Testing for bronchial asthma or chronic obstructive pulmonary disease by
PT comparing the expression level of a marker gene in a biological sample

PT from a subject with the expression level of the gene in a sample from a
PT healthy subject.
PS Claim 1; SEQ ID NO 156; 241pp; English.
XX
CC The present invention describes a method of testing for bronchial asthma
CC or chronic obstructive pulmonary disease. The method comprises
CC determining the expression level of a marker gene in a biological sample
CC from a subject, comparing the expression level determined with the
CC expression level of the marker gene in a biological sample from a healthy
CC subject, and judging whether the subject has bronchial asthma or chronic
CC obstructive pulmonary disease. The marker gene comprises: (a) a group of
CC genes (S1) whose expression levels increase when respiratory epithelial
CC cells are stimulated with interleukin-13; or (b) a group of genes (S2)
CC whose expression levels decrease when respiratory epithelial cells are
CC stimulated with interleukin-13. Also described: (1) a reagent (I) for
CC testing for bronchial asthma or chronic obstructive pulmonary disease;
CC (2) a kit for screening for a candidate compound for a therapeutic agent
CC to treat bronchial asthma or chronic obstructive pulmonary disease; (3)
CC an animal model for bronchial asthma or chronic obstructive pulmonary
CC disease; (4) an inducer that induces bronchial asthma in a mouse; (5) a
CC method for producing an animal model for bronchial asthma or chronic
CC obstructive pulmonary disease; (6) a therapeutic agent for bronchial
CC asthma or chronic obstructive pulmonary disease, comprising the compound,
CC a marker gene or an antisense nucleic acid corresponding to a portion of
CC the marker gene, a ribozyme, a polynucleotide that suppresses the
CC expression of the gene through an RNAi effect or an antibody recognising
CC a protein encoded by a marker gene; and (7) a DNA chip for testing for
CC bronchial asthma or a chronic obstructive pulmonary disease, on which a
CC probe has been immobilised to assay a marker gene. (I) has respiratory
CC and antiasthmatic activities, and can be used in gene therapy. The method
CC is useful for testing for or screening for a therapeutic agent for
CC bronchial asthma or chronic obstructive pulmonary disease. The present
CC sequence is used in the exemplification of the present invention.
XX
SQ Sequence 3380 BP; 1153 A; 644 C; 753 G; 830 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 0 Length: 3380
Score: 5311.00 Matches: 1025
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

US-09-515-363C-2 (1-1025) x ADJ74904 (1-3380)
QY 1 MetSerAsnGlyTyrSerThrAspGluAsnPheArgTyrLeuIleSerCysPheArgAla 20
DB 169 ATGTCGAATGGGTATTCACAGACGAGAAATTCGCTATCTCATCTCGTGTTCAGGGCC 228
QY 21 ArgValLysMetTyrIleGlnValGluProValLeuAspTyrLeuThrPheLeuProAla 40
DB 229 AGGGTGAAATGTATACATCCAGGTGGAGCCTGTGCTGGACTACCTGACCTTTCTGCCTGCA 288
QY 41 GluValLysGluGlnIleGlnArgThrValAlaThrSerGlyAsnMetGlnAlaValGlu 60
DB 289 GAGGTGAAGGAGCGAGATTCAGAGGACAGTCGCCACCTCCGGGAACATGCAGGCAGTTGAA 348
QY 61 LeuLeuLeuSerThrLeuGluLysGlyValTrpHisLeuGlyTrpThrArgGluPheVal 80
DB 349 CTGCTGCTGAGCACCTTGGAGAAGGGAGTCTGGCACCTTGGTTGGACTCGGGAATTCGTG 408
QY 81 GluAlaLeuArgArgThrGlySerProLeuAlaAlaArgTyrMetAsnProGluLeuThr 100
DB 409 GAGGCCCTCCGGAGAACCGGCGAGCCCTCTGGCCCGCGCTACATGAACCTGAGCTCAGC 468
QY 101 AspLeuProSerProSerPheGluAsnAlaHisAspGluTyrLeuGlnLeuLeuAsnLeu 120
DB 469 GACTTGCCCTCTCCATCGTTTGAGAACGCTCATGATGAATATCTCCAACTGCTGAACCTC 528
QY 121 LeuGlnProThrLeuValAspLysLeuLeuValArgAspValLeuAspLysCysMetGlu 140

Db 529 CTTACGCCCACTCTGGTGGACAAGCTTCTTAGTTAGAGACGCTTTGGATAAGTGCATGGAG 588
QY 141 GluGluLeuLeuThrIleGluAspArgAsnArgIleAlaAlaGluAsnAsnGlyAsn 160
Db 589 GAGGAACCTGTTGACAAATTGAAGACAGAAACCGGATTGCTGCTGCAGAAAACAATGGAAT 648
QY 161 GluSerGlyValArgGluLeuLeuLysArgIleValGlnLysGluAsnTrpPheSerAla 180
Db 649 GAATCAGGTGTAAGAGAGCTACTAAAAGGATTGTGCAGAAAGAAAACCTGTTCTCTGCA 708
QY 181 PheLeuAsnValLeuArgGlnThrGlyAsnAsnGluLeuValGlnGluLeuThrGlySer 200
Db 709 TTTCTGAATGTTCTTCTCGTCAACAGGAAACAATGAACCTTGTCCAAGAGTTAACAGGCTCT 768
QY 201 AspCysSerGluSerAsnAlaGluIleGluAsnLeuSerGlnValAspGlyProGlnVal 220
Db 769 GATTGCTCAGAAAAGCAATGCAGAGATTGAGAATTTATCAAGTTGATGGTCTCTCAAGTG 828
QY 221 GluGluGlnLeuLeuSerThrThrValGlnProAsnLeuGluLysGluValTrpGlyMet 240
Db 829 GAAGAGCAACTTCTTCAACCACAGATTGAGCAATTTATCAAGTTGATGGTCTCTCAAGTG 888
QY 241 GluAsnAsnSerGluSerSerPheAlaAspSerValValSerGluSerAspThr 260
Db 889 GAGATAACTCATCAGAAATCATCTTTGAGATTCTTCTGTAGTTTTCAGAAATCAGACACA 948
QY 261 SerLeuAlaGluGlySerValSerCysLeuAspGluSerLeuGlyHisAsnSerAsnMet 280
Db 949 AGTTTGGCAGAGGAAGTGTGAGTCTGCTTAGATGAAAGTCTTGACATACAGCAACATG 1008
QY 281 GlySerAspSerGlyThrMetGlySerAspSerAspGluGluAsnValAlaAlaArgAla 300
Db 1009 GGCAGTGATTGAGCACCATGGGAAGTGTATTCAGATGAAGAGAAATGTGGCAGCAAGACA 1068
QY 301 SerProGluProGluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaLeu 320
Db 1069 TCCCCGGAGCCAGAACTCCAGCTCAGGCCTTACCAATGGAAGTTGCCAGCCAGCCTTG 1128
QY 321 GluGlyLysAsnIleIleIleCysLeuProThrGlySerGlyLysThrArgValAlaVal 340
Db 1129 GAAGGGAAGAAATATCATCATCTGCCTCCCTACAGGGAGTGGAAAACACAGAGTGGCTGT 1188
QY 341 TyrIleAlaLysAspHisLeuAspLysLysLysAlaSerGluProGlyLysValIle 360
Db 1189 TACATTGCCAAGGATCACTTAGACAAGAAAGAAAAGCATCTGAGCCTGGAAAAGTTATA 1248
QY 361 ValLeuValAsnLysValLeuLeuValGluGlnLeuPheArgLysGluPheGlnProPhe 380
Db 1249 GTTCTTGTCATAAAGTACTGTCTAGTTGAACAGCTCTTCCGCAAGGAGTTCCAAACCATTT 1308
QY 381 LeuLysLysTrpTyrArgValIleGlyLeuSerGlyAspThrGlnLeuLysIleSerPhe 400
Db 1309 TTGAAGAAATGGTATCGTGTATTATGGATTAAAGTGGTGATACCCAACTGAAAATATCATTT 1368
QY 401 ProGluValValLysSerCysAspIleIleIleSerThrAlaGlnIleLeuGluAsnSer 420
Db 1369 CCAGAGTTGTCAAGTCTCTGTGATATTTATTCAGTACAGCTCAATTCCTTCAAACTCC 1428
QY 421 LeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPheSerLeuIle 440
Db 1429 CTCITAAACTTGGAAAATGGAGAGATGCTGGTGTTCATTTGTGAGACTTTTCCCTCAT 1488
QY 441 IleIleAspGluCysHisThrAsnLysGluAlaValTyrAsnAsnIleMetArgHis 460
Db 1489 ATCAATTGATGAATGTTCATCACCAACAAGAGCAGTGTATATAATACATCATGAGGCAT 1548
QY 461 TyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysLysGluAsnLysProValIlePro 480
Db 1549 TATTGATGCAGAAAGTTGAAAACAATAGACTCAGAAAGAAAACAACCAAGTATTTCCC 1608
QY 481 LeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyGlyAlaThrLysGlnAla 500
Db 1609 CTTCTCATGATCTGGGACTAACAGCTTCACTGCTGGTGTGGAGGGGGCCCAAGCAAGCC 1668

QY 501 LysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThrIleLysThr 520
Db 1669 AAAGCTGAAGAACACATTTTAAACTATGTGCCAATCTTGATGCATTTACTATTAAACT 1728

QY 521 ValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluProCysLysLysPheAla 540
Db 1729 GTTAAAGAAAACCTTGATCAACTGAAAAACCAATACAGGAGCCATGCAAGAGTTTGCC 1788

QY 541 IleAlaAspAlaThrArgGluAspProPheLysGluLysLeuLeuGluIleMetThrArg 560
Db 1789 ATTGCAGATGCACAGAGAAGATCCATTTAAAGAGAAACTTCTAGAAATAATGACAAGG 1848

QY 561 IleGlnThrTyrCysGlnMetSerProMetSerAspPheGlyThrGlnProTyrGluGln 580
Db 1849 ATTCAAACTTATTGTCAAATGAGTCCCAATGTCCAGATTTTGGAACTCAACCCCTATGAACAA 1908

QY 581 TrpAlaIleGlnMetGluLysLysAlaAlaLysLysGlyAsnArgLysGluArgValCys 600
Db 1909 TGGGCCATTCAATGGAAAAAAGCTGCAAAAAAAGGAAATCGCAAGAACGTTGTTGT 1968

QY 601 AlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMetIle 620
Db 1969 GCAGAACATTTGAGGAAGTACAATGAGGCCCTACAAATTATGACACAAATTCGAATGATA 2028

QY 621 AspAlaTyrThrHisLeuGluThrPheTyrAsnGluGluLysAspLysLysPheAlaVal 640
Db 2029 GATCGGTATACTCATCTTGAAACTTTCTATAATGAAGAGAAAGATAAGAAGTTTGCAGTC 2088

QY 641 IleGluAspSerAspGluGlyGlyAspAspGluTyrCysAspGlyAspGluAspGlu 660
Db 2089 ATAGAAGATGATAGTGATGAGGGTGGTGATGATGAGTATTGTGATGGTGATGAAGATGAG 2148

QY 661 AspAspLeuLysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeuphe 680
Db 2149 GATGATTTAAAGAAAACCTTTGAAACTGGATGAAACAGATAGATTCTCATGACTTTATTT 2208

QY 681 PheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGluLys 700
Db 2209 TTTGAAAAACAATAAAATGTTGAAAAGGTGGCTGAAACCCCAATATGAAAATGAAAAG 2268

QY 701 LeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArgThrGluGluSerAlaArg 720
Db 2269 CTGACCAAATTAAGAAATACCATAATGGAGCAATATACTAGGACTGAGGAATCAGCACGA 2328

QY 721 GlyIleIlePheThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTrpIleThrGlu 740
Db 2329 GGAATAATCTTTACAAAAAACACGACAGAGTGCAATATGCGCTTTCCAGTGGATTACTGAA 2388

QY 741 AsnGluLysPheAlaGluValGlyValLysAlaHisHisLeuIleGlyAlaGlyHisSer 760
Db 2389 AATGAAAAATTTGCTGAAGTAGGAGTCAAAGCCCACCATCTGATTGGAGCTGGACACAGC 2448

QY 761 SerGluPheLysProMetThrGlnAsnGluGlnLysGluValIleSerLysPheArgThr 780
Db 2449 AGTGAGTTCAAACCCCATGACACAGAAATGAACAAAAGAAGTATTAGTAAATTTCCGACT 2508

QY 781 GlyLysIleAsnLeuIleAlaThrThrValAlaGluGluGlyLeuAspIleLysGlu 800
Db 2509 GGAAAAATCAATCTGCTTATCGCTACCACAGTGGCAGAGAAGAGGTCTGGATATTAAAGAA 2568

QY 801 CysAsnIleValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAlaArg 820
Db 2569 TGTAACATTGTTATCCGTTATGGTCTCGTCACCAATGAAATAGCCATGGTCCAGGCCGT 2628

QY 821 GlyArgAlaArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGlyVal 840
Db 2629 GGTGAGCCAGAGCTGATGAGAGCACCTACGTCTGTTGCTCACAGTGGTTCAGGAGTT 2688

QY 841 IleGluHisGluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHisCys 860
Db 2689 ATCGAACATGAGACAGTTAATGATTTCCGAGAGAAGATGATGATATATAAGCTATACATTGT 2748

QY 861 ValGlnAsnMetLysProGluGluTyrAlaHisLysIleLeuGluLeuGlnMetGlnSer 880
Db 2749 GTTCAAAATATGAAACCAGAGGAGTATGCTCATAAGATTTTGGAAATTACAGATGCAAGT 2808

QY 881 IleMetGluLysLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAsnPro 900
Db 2809 ATAATGGAAGAAAATGAAACCAAGAGAAATATTGCCAAGCATTACAAGAAATAACCCA 2868

QY 901 SerLeuIleThrPheLeuCysLysAsnCysSerValLeuAlaCysSerGlyGluAspIle 920
Db 2869 TCACATAATACTTTCCTTTGCAAAAACTGCAGTGTCTAGCTGTTCTGGGAAGATATC 2928

QY 921 HisValIleGluLysMetHisHisValAsnMetThrProGluPheLysGluLeuTyrIle 940
Db 2929 CATGTAATTGAGAAAATGCATCACGTCAATATGACCCAGAAATTCAGGAACCTTTACATT 2988

QY 941 ValArgGluAsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluIle 960
Db 2989 GTAAGAGAAAAACAAAGCACTGCAAAAAGAAAGTGTCCGACTATCAAAATAAATGGTGAATC 3048

QY 961 IleCysLysCysGlyGlnAlaTrpGlyThrMetMetValHisLysGlyLeuAspLeuPro 980
Db 3049 ATCTGCAAAATGTGGCCAGGCTTGGGAAACAATGATGGTGCAAAAAGGCTTAGATTTCCT 3108

QY 981 CysLeuLysIleArgAsnPheValValPheLysAsnAsnSerThrLysLysGlnTyr 1000
Db 3109 TGTCCTCAAAATAAGGAATTTTGTAGTGGTTTTCAAAAAATAATTCAACAAAGAAACAATAC 3168

QY 1001 LysLysTrpValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCysLeu 1020
Db 3169 AAAAGTGGTAGAATTACCTATCACATTTCCCAATCTTGACTATTTCAGAAATGCTGTTTA 3228

QY 1021 PheSerAspGluAsp 1025
Db 3229 TTTAGTGATGAGGAT 3243

RESULT 3
ABA04908
ID ABA04908 standard; cDNA; 3372 BP.
XX
AC ABA04908;
XX
DT 05-MAR-2002 (first entry)
XX
DE Human RNA helicase RH116 coding sequence.
XX
KW Human; RH116; RNA helicase; cytostatic; virucide; anti-HIV;
KW immunosuppressive; immunostimulatory; antirheumatic; antiarthritic;
KW antiarteriosclerotic; osteopathic; antidiabetic; hepatotropic;
KW antiinflammatory; cancer; infection; HIV; hepatitis; genetic disease;
KW autoimmune disease; graft rejection; vaccine; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 155..3232
FT /*tag= a
FT /product= "Human RH116"
XX
PN WO200185955-A1.
XX
PD 15-NOV-2001.
XX
PF 11-MAY-2001; 2001WO-FR001441.
XX
PR 11-MAY-2000; 2000FR-00006030.
XX
PA (ISTA-) ISTAC.
PA (INSP) INST PASTEUR LILLE.
XX
PI Bahr G, Cocude C, Capron A;
XX
DR WPI; 2002-082898/11.

DR P-PSDB; AAM47798.
XX
PT New polypeptide, useful for treating and diagnosing cancer or
PT inflammation, and drug screening, comprises a human polynucleotide
PT homologous to RNA helicase.
XX
PS Claim 7; Page 85-89; 114pp; French.
XX
CC The present sequence is the coding sequence for human RH116. RH116 is a
CC 116kDa protein and has homology to RNA helicases (DEXH box). RH116 and
CC its coding sequence are useful for treating cancer; acute or chronic
CC infections (especially by HIV or hepatitis B or C); inherited genetic
CC diseases; (auto)immune diseases (particularly rheumatism, arthritis,
CC arteriosclerosis, osteoporosis and diabetes, but many others listed) and
CC to prevent graft rejection. RH116 and its coding sequence are also useful
CC for inducing, or increasing, the immune response to a vaccine
XX
SQ Sequence 3372 BP; 1157 A; 637 C; 748 G; 830 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 0 Length: 3372
Score: 5285.00 Matches: 1020
Percent Similarity: 99.71% Conservative: 2
Best Local Similarity: 99.51% Mismatches: 3
Query Match: 99.51% Indels: 0
DB: 6 Gaps: 0

US-09-515-363C-2 (1-1025) x ABA04908 (1-3372)

QY 1 MetSerAsnGlyTyrSerThrThrAspGluAsnPheArgTyrLeuIleSerCysPheArgAla 20
Db
QY 21 ArgValLysMetTyrIleGlnValGluProValLeuAspTyrLeuThrPheLeuProAla 40
Db
QY 41 GluValLysGluGlnIleGlnArgThrValAlaThrSerGlyAsnMetGlnAlaValGlu 60
Db
QY 61 LeuLeuLeuSerThrLeuGluLysGlyValTrpHisLeuGlyTrpThrArgGluPheVal 80
Db
QY 81 GluAlaLeuArgArgThrGlySerProLeuAlaAlaArgTyrMetAsnProGluLeuThr 100
Db
QY 101 AspLeuProSerProSerPheGluAsnAlaHisAspGluTyrLeuGlnLeuLeuAsnLeu 120
Db
QY 121 LeuGlnProThrLeuValAspLysLeuLeuValArgAspValLeuAspLysCysMetGlu 140
Db
QY 141 GluGluLeuLeuThrIleGluAspArgAsnArgIleAlaAlaGluAsnAsnGlyAsn 160
Db
QY 161 GluSerGlyValArgGluLeuLeuLysArgIleValGlnLysGluAsnTrpPheSerAla 180
Db
QY 181 PheLeuAsnValLeuArgGlnThrGlyAsnAsnGluLeuValGlnGluLeuThrGlySer 200
Db
QY 201 AspCysSerGluSerAsnAlaGluIleGluAsnLeuSerGlnValAspGlyProGlnVal 220
Db
QY 241 GluAsnAsnSerSerGluSerSerPheAlaAspSerSerValValSerGluSerAspThr 260
Db
QY 281 GlySerAspSerGlyThrMetGlySerAspSerAspGluGluAsnValAlaAlaArgAla 300
Db
QY 301 SerProGluProGluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaLeu 320
Db
QY 321 GluGlyLysAsnIleIleCysLeuProThrGlySerGlyLysThrArgValAlaVal 340
Db
QY 341 TyrIleAlaLysAspHisLeuAspLysLysLysAlaSerGluProGlyLysValIle 360
Db
QY 361 ValLeuValAsnLysValLeuLeuValGluGlnLeuPheArgLysGluPheGlnProphe 380
Db
QY 381 LeuLysLysTrpTyrArgValIleGlyLeuSerGlyAspThrGlnLeuLysIleSerPhe 400
Db
QY 401 ProGluValValLysSerCysAspIleIleSerThrAlaGlnIleLeuGluAsnSer 420
Db
QY 421 LeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPheSerLeuIle 440
Db
QY 441 IleIleAspGluCysHisHisThrAsnLysGluAlaValTyrAsnAsnIleMetArgHis 460
Db
QY 461 TyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysLysGluAsnLysProValIlePro 480
Db
QY 481 LeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyAlaThrLysGlnAla 500
Db
QY 501 LysAlaGluGluHisIleLeuLysLeuCysAlaAsnGlnIleGlnGluProCysLysLysPheAla 540
Db
QY 521 VallysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluProCysLysLysPheAla 560
Db
QY 541 IleAlaAspAlaThrArgGluAspPropheLysGluLysLeuLeuIleMetThrArg 580
Db
QY 561 IleGlnThrTyrCysGlnMetSerProMetSerAspPheGlyThrGlnProTyrGluGln 580
Db
QY 581 TrpAlaIleGlnMetGluLysLysAlaAlaLysLysGlyAsnArgLysGluArgValCys 600

221 GluGluGlnLeuLeuSerThrThrValGlnProAsnLeuGluLysGluValTrpGlyMet 240
Db
815 GAAGAGCAACTTCTTTCAACCAACAGTTCAGCCAAATCTGGAGAAGGAGGTCTGGGGCATG 874
QY
241 GluAsnAsnSerSerGluSerSerPheAlaAspSerSerValValSerGluSerAspThr 260
Db
875 GAGAATAACTCATCAGAATCATCTTTTGCAGATTCTTCTGTAGTTTCAGAATCAGACACA 934
QY
261 SerLeuAlaGluGlySerValSerCysLeuAspGluSerLeuGlyHisAsnSerAsnMet 280
Db
935 AGTTTGGCAGAGAAGAGTGTACAGTCTGTAGATGAAAGTCTTGGACATAAACAGCAACATG 994
QY
281 GlySerAspSerGlyThrMetGlySerAspSerAspGluGluAsnValAlaAlaArgAla 300
Db
995 GGCAGTGATTCAGGCACCATGGGAAGTGATTTCAGATGAAGAGAATGTGGCAGCAAGAGCA 1054
QY
301 SerProGluProGluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaLeu 320
Db
1055 TCCCCGGAGCCAGAACTCCAGCTCAGGCCTTACCAATGGAAATGGAGTTGCCAGCCCTTG 1114
QY
321 GluGlyLysAsnIleIleCysLeuProThrGlySerGlyLysThrArgValAlaVal 340
Db
1115 GAAGGGAAGAATATCATCTGCCTCCCTACAGGGAGTGGAAAAACACAGATGGCTGTT 1174
QY
341 TyrIleAlaLysAspHisLeuAspLysLysLysAlaSerGluProGlyLysValIle 360
Db
1175 TACATTGCCAAGGATCACTTAGACAAGAAAGAAAGCATCTGAGCCTGGAAAAAGTTATA 1234
QY
361 ValLeuValAsnLysValLeuLeuValGluGlnLeuPheArgLysGluPheGlnProphe 380
Db
1235 GTTCTTGTCAATAAGGTACTGTCTAGTTGAACAGCTCTTCCGCAAGAGATTCCAAACCATTT 1294
QY
381 LeuLysLysTrpTyrArgValIleGlyLeuSerGlyAspThrGlnLeuLysIleSerPhe 400
Db
1295 TTGAAGAAATGGTATCGTGTATTATGGATTAAAGTGGTGATACCCAACTGAAAAATATCATTT 1354
QY
401 ProGluValValLysSerCysAspIleIleSerThrAlaGlnIleLeuGluAsnSer 420
Db
1355 CCAGAAAGTTGTCAAGTCTGTGATATTATTATCAGTACAGTCAATCTTGAAGAACTCC 1414
QY
421 LeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPheSerLeuIle 440
Db
1415 CTCTTAAACTTGGAAAAATGGAGAAGATGCTGGTGTTCATTGTGCAGACTTTCTCTTCATT 1474
QY
441 IleIleAspGluCysHisHisThrAsnLysGluAlaValTyrAsnAsnIleMetArgHis 460
Db
1475 ATCATTTGATGAATGTATCATCACCAACAAGAGCAGTGTATATAATACATCATGAGGCAT 1534
QY
461 TyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysLysGluAsnLysProValIlePro 480
Db
1535 TATTTGATGCAGAAAGTTGAAAAACAATAGACTCAAGAAAGAAAAACAACACAGTATCCC 1594
QY
481 LeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyAlaThrLysGlnAla 500
Db
1595 CTTCTCTCAGATACTGGGACTAAACAGCTTACCTGGTGTGGAGGGGCCACGAAGCAAGCC 1654
QY
501 LysAlaGluGluHisIleLeuLysLeuCysAlaAsnGlnIleGlnGluProCysLysLysPheAla 540
Db
1655 AAAGCTGAAGAACACATTTTAAAAACTATGTGCAATCTTGTATGCATTTACTATAAACT 1714
QY
521 VallysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluProCysLysLysPheAla 560
Db
1715 GTTAAAGAAAAACCTTTGATCAACTGAAAAACCAATAACAGGAGCCATGCAAGAGTTTGCC 1774
QY
541 IleAlaAspAlaThrArgGluAspPropheLysGluLysLeuLeuIleMetThrArg 580
Db
1775 ATTGCAGATGCAACACAGAGAAGATCCATTTTAAAGAGAAAACTTCTAGAAAATAATGACAAGG 1834
QY
561 IleGlnThrTyrCysGlnMetSerProMetSerAspPheGlyThrGlnProTyrGluGln 580
Db
1835 ATTCAAACTTATTGTCAATGAGTCCCAATGTGAGATTTTGGAACTCAACCTATGAACAA 1894
QY
581 TrpAlaIleGlnMetGluLysLysAlaAlaLysLysGlyAsnArgLysGluArgValCys 600

Db 1895 TGGGCCATTCAATGGAAAAAAGCTGCAAAAGAAAGAAATCGCAAAAGAAAGTGTGTGT 1954
QY 601 AlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMetIle 620
Db 1955 GCAGAACATTTGAGGAAGTACAATAAGGCCCTACAAATTAATGACACAATTCGAATGATA 2014
QY 621 AspAlaTyrThrHisLeuGluThrPheTyrAsnGluGluLysAspLysLysPheAlaVal 640
Db 2015 GATCGGTATACTCATCTTGAAACCTTCTATAATGAAGAGAAAGATAAGAAGTTTGCACTC 2074
QY 641 IleGluAspAspSerAspGluGlyGlyAspAspGluTyrCysAspGlyAspGluAspGlu 660
Db 2075 ATAGAAGATGATAGTATGAGGGTGGTGATGATGAGTATTGTGATGGTGATGAAGATGAG 2134
QY 661 AspAspLeuLysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeuPhe 680
Db 2135 GATGATTTAAAGAAACCTTTGAAACTGGATGAAACAGATAGATTTCTCATGACTTTATTT 2194
QY 681 PheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGluLys 700
Db 2195 TTTGAAAAACAATAAAATGTTGAAAAGGCTGGCTGAAAACCCAGAATATGAAAATGAAAAG 2254
QY 701 LeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArgThrGluGluSerAlaArg 720
Db 2255 CTGACCAAAATTAAGAAATACCATAATGGAGCAATATACTAGGACTGAGGAATCAGCACGA 2314
QY 721 GlyIleIlePheThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTrpIleThrGlu 740
Db 2315 GGAATAATCTTTACAAAAACACGACAGAGTGCATATGCGCTTTCCCAGTGGATTACTGAA 2374
QY 741 AsnGluLysPheAlaGluValGlyValLysAlaHisHisLeuIleGlyAlaGlyHisSer 760
Db 2375 AATGAAAAATTTGCTGAAGTAGGAGTCAAAGCCACCATCTGATTGGAGCTGGACACAGC 2434
QY 761 SerGluPheLysProMetThrGlnAsnGluGlnLysGluValIleSerLysPheArgThr 780
Db 2435 AGTGAGTTCAAAACCCATGACACAGAAATGAACAAAAGAAAGTCATTAGTAAATTTCCGACT 2494
QY 781 GlyLysIleAsnLeuLeuIleAlaThrThrValAlaGluGluGlyLeuAspIleLysGlu 800
Db 2495 GGAAAAATAATCTGCTTATCGCTACACAGTGGCAGAGAGAGGTCTGGATATTAAGAA 2554
QY 801 CysAsnIleValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAlaArg 820
Db 2555 TGTAACATTGTTATCCGTTATGGTCTCGTCACCAATGAAATAGCCATGGTCCAGGCCGT 2614
QY 821 GlyArgAlaArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGlyVal 840
Db 2615 GGTCGAGCCAGAGCTGATGAGAGCACCTACGTCCTGGTTGCTCACAGTGGTTCAGAGTT 2674
QY 841 IleGluHisGluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHisCys 860
Db 2675 ATCGAACGTGAGACAGTTAATGATTTCCGAGAGAAGATGATGTATATAAGCTATACATTGT 2734
QY 861 ValGlnAsnMetLysProGluGluTyrAlaHisLysIleLeuGluLeuGlnMetGlnSer 880
Db 2735 GTTCAAAAATATGAACCCAGAGGAGTATGCTCATAAGATTTTGAATTACAGATGCAAGT 2794
QY 881 IleMetGluLysLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAsnPro 900
Db 2795 ATAATGGAAAAGAAAATGAAAACCAAGAGAAAATATTGCCACGATTACAAGAAATAACCCA 2854
QY 901 SerLeuIleThrPheLeuCysLysAsnCysSerValLeuAlaCysSerGlyGluAspIle 920
Db 2855 TCACATAATACTTCTCTTTGCAAAAACCTGCAGTGTGTAGCCTGTTCTGGGAAGATATC 2914
QY 921 HisValIleGluLysMetHisHisValAsnMetThrProGluPheLysGluLeuTyrIle 940
Db 2915 CATGTAATTGAGAAAAATGCATCACGTCAAATATGACCCCAAGAAATCAAGGAACCTTTACATT 2974
QY 941 ValArgGluAsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluIle 960

Db 2975 GTAAGAGAAAAACAAGCACTGCAAAAGAAAGTGTCCGACTATCAAAATAATGGTGAATC 3034
QY 961 IleCysLysCysGlyGlnAlaTrpGlyThrMetMetValHisLysGlyLeuAspLeuPro 980
Db 3035 ATCTGCAAATGTGCCAGGCTTGGGAAACAATGATGGTGCAAAAAGGCTTAGATTTCCT 3094
QY 981 CysLeuLysIleArgAsnPheValValPheLysAsnAsnSerThrLysLysGlnTyr 1000
Db 3095 TGTCTCAAAATAAGGAATTTTGTAGTGGTTTTCAAAAATAATTCAACAAAGAAACAATAC 3154
QY 1001 LysLysTrpValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCysCysLeu 1020
Db 3155 AAAAAGTGGTAGAATTACCTATCACATTTCCCAATCTTGACTATTTCAGAATGCTGTTA 3214
QY 1021 PheSerAspGluAsp 1025
Db 3215 TTTAGTGATGAGGAT 3229
RESULT 4
ADC30823
ID ADC30823 standard; cDNA; 3446 BP.
XX
AC ADC30823;
XX
DT 18-DEC-2003 (first entry)
XX
DE Human novel cDNA sequence, SEQ ID NO:905.
XX
KW Human; diagnostic; drug screening; forensics; gene mapping;
KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;
KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
KW ulcers; osteoporosis; autoimmune disease; cancer;
KW molecular weight marker; food supplement; antiparkinsonian; nootropic;
KW neuroprotective; antianaemic; anticoagulant; thrombolytic; vulnerary;
KW antiulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
KW gene therapy; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2003029271-A2.
XX
PD 10-APR-2003.
XX
PF 24-SEP-2002; 2002WO-US030474.
XX
PR 24-SEP-2001; 2001US-0324631P.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;
PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
PI Haley-Vicente D, Drmanac RT;
XX
DR WPI; 2003-371981/35.
DR P-PSDB; ADC31794.
XX
PT New polynucleotide and polypeptide useful for diagnosing, preventing or
PT treating conditions such as neurodegenerative diseases, anemias, platelet
PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
PT cancer.
XX
PS Claim 1; SEQ ID NO 905; 1185pp; English.
XX
CC The invention relates to 971 novel human cDNA sequences (ADC29919-
CC ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The
CC invention also relates to nucleic acid sequences over 99% identical with
CC the novel human cDNAs. The invention additionally encompasses expression
CC vectors and host cells comprising a nucleic acid of the invention; the
CC recombinant production of a polypeptide of the invention; an antibody
CC against a polypeptide of the invention; a method of detecting
CC polynucleotides or polypeptides of the invention; and methods of
CC identifying a compound which binds to a polypeptide of the invention. The
CC invention further discloses methods of preventing, treating or

CC ameliorating a medical condition; kits comprising polynucleotide probes
CC and/or monoclonal antibodies for carrying out the methods of the
CC invention; methods for the identification of compounds that modulate the
CC expression or activity of the polynucleotide and/or polypeptide; and 767
CC contig sequences corresponding to the cDNA sequences of the invention
CC (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628
CC -ADC33394). The nucleic acids and polypeptides of the invention are
CC useful in diagnostics, drug screening, forensics, gene mapping, in the
CC identification of mutations responsible for genetic disorders or other
CC traits, for assessing biodiversity, and in producing many other types of
CC data and products dependent on DNA and amino acid sequences. They are
CC also used for treating diseases such as Parkinson's disease, Alzheimer's
CC disease and other neurodegenerative diseases, anaemia, platelet
CC disorders, wounds, ulcers, osteoporosis, autoimmune diseases or
CC cancer. The nucleic acids may also be used as hybridisation probes or
CC primers, and in the recombinant production of a protein. The polypeptides
CC are also useful in generating antibodies, as molecular weight markers,
CC and as food supplements. The present sequence represents a specifically
CC claimed human cDNA sequence of the invention. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 3446 BP; 1184 A; 649 C; 772 G; 841 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 0 Length: 3446
Score: 5285.00 Matches: 1020
Percent Similarity: 99.71% Conservative: 2
Best Local Similarity: 99.51% Mismatches: 3
Query Match: 99.51% Indels: 0
DB: 10 Gaps: 0

US-09-515-363C-2 (1-1025) x ADC30823 (1-3446)

QY	1	MetSerAsnGlyTyrSerThrAspGluAsnPheArgTyrLeuIleSerCysPheArgAla	20
DB	237	ATGTCGAATGGGTATTCACAGACGAGAATTTCCGCTATCTCATCTCGTGTTCAGGGCC	296
QY	21	ArgValLysMetTyrIleGlnValGluProValLeuAspTyrLeuThrPheLeuProAla	40
DB	297	AGGGTGAAATGTACATCCAGGTGGAGCCTGTGTGGACTACCTGACCTTTCTGCCTGCA	356
QY	41	GluValLysGluGlnIleGlnArgThrValAlaThrSerGlyAsnMetGlnAlaValGlu	60
DB	357	GAGGTGAAGGAGCAGATTCCAGAGGACATCGCCACCTCCGGAAACATGCAGGCAGTTGAA	416
QY	61	LeuLeuLeuSerThrLeuGluLysGlyValTrpHisLeuGlyTrpThrArgGluPheVal	80
DB	417	CTGCTGCTGAGCACCTTGAGAGGGAGTCTGGCACCTTGGTGGACTCGGGAATTCGTG	476
QY	81	GluAlaLeuArgThrGlySerProLeuAlaAlaArgTyrMetAsnProGluLeuThr	100
DB	477	GAGGCCCTCCGGAGAACCGGCAGCCCTCTGGCCGCCCGCTACATGAACCTTGAGTCAAG	536
QY	101	AspLeuProSerProSerPheGluAsnAlaHisAspGluTyrLeuGlnLeuLeuAsnLeu	120
DB	537	GACTTGCCCTCTCCATCGTTTGAGAACGCTCATGATGAATATCTCCAACCTGCTGAACCTC	596
QY	121	LeuGlnProThrLeuValAspLysLeuLeuValArgAspValLeuAspLysCysMetGlu	140
DB	597	CTTCAGCCCACTCTGCTGGACAAGCTTCTAGTTAGAGACGCTCTTGATAAGTGCATGGAG	656
QY	141	GluGluLeuLeuThrIleGluAspArgAsnArgIleAlaAlaGluAsnAsnGlyAsn	160
DB	657	GAGGAACCTGTTGACAAATGAAGACAGAAACCGGATTGCTGTCAGAAAAACAATGGAAAT	716
QY	161	GluSerGlyValArgGluLeuLeuLysArgIleValGlnLysGluAsnTrpPheSerAla	180
DB	717	GAATCAGGTGTAAGAGAGCTACTAAAAAGGATTGTGCAGAAAGAAAACTGGTTCTCTGCA	776
QY	181	PheLeuAsnValLeuArgGlnThrGlyAsnAsnGluLeuValGlnLeuThrGlySer	200

DB	777	TTTCTGAATGTTCTTCGTCAAAACAGGAAAAAACAATGAACCTGTCCAAGAGTTAACAGGCTCT	836
QY	201	AspCysSerGluSerAsnAlaGluIleGluAsnLeuSerGlnValAspGlyProGlnVal	220
DB	837	GATTGCTCAGAAAGCAATGCAGAGATTGAGAAATTTATCACAAGTTGATGGTCTCTCAAGTG	896
QY	221	GluGluGlnLeuLeuSerThrThrValGlnProAsnLeuGluLysGluValTrpGlyMet	240
DB	897	GAAGAGCAACTTCTTTCAACCACAGTTCAGCCAAATCTGGAGAAGGAGGTCTGGGGCATG	956
QY	241	GluAsnAsnSerSerGluSerSerPheAlaAspSerSerValValSerGluSerAspThr	260
DB	957	GAGAAATAACTCATCAGAAATCATCTTTTCAGATTCTTCTGTAGTTTCAGAAATCAGACACA	1016
QY	261	SerLeuAlaGluGlySerValSerCysLeuAspGluSerLeuGlyHisAsnSerAsnMet	280
DB	1017	AGTTTGGCAGAAAGAGTGTCAAGCTGCTTAGATGAAAGTCTTGACATACAGCAACATG	1076
QY	281	GlySerAspSerGlyThrMetGlySerAspSerAspGluGluAsnValAlaAlaArgAla	300
DB	1077	GGCAGTGATTCAGGCACCATGGGAAGTGATTTCAGATGAAGAGAATGTGGCAGCAAGACA	1136
QY	301	SerProGluProGluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaLeu	320
DB	1137	TCCCCGGAGCCAGAACTCCAGCTCAGGCCTTACCRAATGGAAGTTGCCAGCCAGCCTTG	1196
QY	321	GluGlyLysAsnIleIleCysLeuProThrGlySerGlyLysThrArgValAlaVal	340
DB	1197	GAAGGGAAGAATATCATCATCTGCCCTCCCTACAGGGAGTGGAAAAAACAGAGTGGTGT	1256
QY	341	TyrIleAlaLysAspHisLeuAspLysLysLysAlaSerGluProGlyLysValIle	360
DB	1257	TACATTGCCAAGGATCACTTAGACAAGAAGAAAAAGCATCTGAGCCTGGAAAAAGTTATA	1316
QY	361	ValLeuValAsnLysValLeuLeuValGluGlnLeuPheArgLysGluPheGlnProPhe	380
DB	1317	GTTCTTGTCAATAAGGTACTGCTAGTTGAACAGCTCTTCCGCAAGGAGTTCCCAACCATTT	1376
QY	381	LeuLysLysTrpTyrArgValIleGlyLeuSerGlyAspThrGlnLeuLysIleSerPhe	400
DB	1377	TTGAAGAAATGGTATCGTGTATTATGGATTAAAGTGGTGATACCCCAACTGAAATATCATTT	1436
QY	401	ProGluValValLysSerCysAspIleIleSerThrAlaGlnIleLeuGluAsnSer	420
DB	1437	CCAGAAGTTGTCAAGTCTGTGATATTATTAATATCAGTACAGTCAAATCCTTGAAACTCC	1496
QY	421	LeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPheSerLeuIle	440
DB	1497	CTCTTAAACTTGGAAAAATGGAGAAGATGCTGGTGTTCATATTGTGAGACTTTTCTCTCAT	1556
QY	441	IleIleAspGluCysHisHisThrAsnLysGluAlaValTyrAsnAsnIleMetArgHis	460
DB	1557	ATCATTTGATGAATGTCATCACACCAACAAGAAGCAGTGTATATAAATCATCATGAGGCAT	1616
QY	461	TyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysLysGluAsnLysProValIlePro	480
DB	1617	TATTTGATGCAGAAAGTTGAAAAACAATAGACTCAAGAAAGAAAAACAACCAAGTATCCC	1676
QY	481	LeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyGlyAlaThrLysGlnAla	500
DB	1677	CTTCTCAGATACTGGGACTAACAGCTTCACTGGTGTGGAGGGGCCACGAAGCAAGCC	1736
QY	501	LysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThrIleLysThr	520
DB	1737	AAAGCTGAAGAAACACATTTTAAAAACTATGTGCCAATCTTGATGCATTTACTATAAACT	1796
QY	521	ValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluProCysLysLysPheAla	540
DB	1797	GTTAAAGAAAAACCTTGATCAACTGAAAAACCAATAACAGGAGCCATGCAAGAAGTTTGCC	1856
QY	541	IleAlaAspAlaThrArgGluAspProPheLysGluLysLeuLeuGluIleMetThrArg	560
DB	1857	ATTGCAGATGCAACCAGAGAAGATCCATTTTAAAGAGAAAACTTCTAGAAATAATGACAAGG	1916

QY 561 IleGlnThrTyrCysGlnMetSerProMetSerAspPheGlyThrGlnProTyrGluGln 580
Db 1917 ATTCAAACTTATTGTCAAATGAGTCCAAATGTCCAGATTTTGGAACTCAACCCCTATGAACA 1976
QY 581 TrpAlaIleGlnMetGluLysLysAlaAlaLysLysGlyAsnArgLysGluArgValCys 600
Db 1977 TGGGCCATTCAATGGAAAAAAGCTGCAAAAGAAAGGAATCGCAAGAAAGTGTGT 2036
QY 601 AlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMetIle 620
Db 2037 GCAGAACATTTGAGGAAGTACAATAAGGCCCTACAAATTAATGACACAATTCGAATGATA 2096
QY 621 AspAlaTyrThrHisLeuGluThrPheTyrAsnGluLysLysAspLysLysPheAlaVal 640
Db 2097 GATCGTATACTCTTTGAAACTTTCTATAATGAAGAGAAAGATAAAGAGTTTGCAATGATC 2156
QY 641 IleGluAspAspSerAspGluGlyGlyAspAspGluTyrCysAspGlyAspGluAspGlu 660
Db 2157 ATAGAAGATGATAGTATGAGGGTGGTGATGATGAGTATTGTGATGGTGATGAAGATGAG 2216
QY 661 AspAspLeuLysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeuPhe 680
Db 2217 GATGATTTAAAGAAACCTTTGAAACTGGATGAACAGATAGATTCTCATGACTTTATT 2276
QY 681 PheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGluLys 700
Db 2277 TTTGAAAAACAATAAAATGTTGAAAAAGGTGGCTGAAACCCAGAAATATGAAAAAG 2336
QY 701 LeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArgThrGluGluSerAlaArg 720
Db 2337 CTGACCAATTAAGAAATACCATAATGAGCAATATACTAGGACTGAGGAATCAGCACGA 2396
QY 721 GlyIleIlePheThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTrpIleThrGlu 740
Db 2397 GGAATAATCTTTACAAAAACACGACAGACTGCATATGCGCTTCCCAGTGGATTACTGAA 2456
QY 741 AsnGluLysPheAlaGluValGlyValLysAlaHisLeuIleGlyAlaGlyHisSer 760
Db 2457 AATGAAAAATTGCTGAAGTAGGAGTCAAGCCCCACCATCTGATTGGAGCTGGACACAGC 2516
QY 761 SerGluPheLysProMetThrGlnAsnGluGlnLysGluValIleSerLysPheArgThr 780
Db 2517 AGTGAGTTCAAAACCCATGACACAGATGAACAAAAAGAGTATTAGTAAATTTCCGACT 2576
QY 781 GlyLysIleAsnLeuLeuIleAlaThrValAlaGluGluGlyLeuAspIleLysGlu 800
Db 2577 GGAAAAATAAATCTGCTTATCGCTACCAAGTGGCAGAGAGGTCTGGATATTAAAGAA 2636
QY 801 CysAsnIleValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAlaArg 820
Db 2637 TGTAACATTGTTATCCGTTATGGTCTCGTCACCAATGAAATAGCATGGTCCAGGCCGT 2696
QY 821 GlyArgAlaArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGlyVal 840
Db 2697 GGTCGAGCCAGAGCTGATGAGAGCACCTACGTCCTGGTTGCTCACAGTGGTTCAGAGATT 2756
QY 841 IleGluHisGluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHisCys 860
Db 2757 ATCGAACGTGAGACAGTTAATGATTTCCGAGAGAAGATGATGATATAAGCTATACATTGT 2816
QY 861 ValGlnAsnMetLysProGluGluTyrAlaHisLysIleLeuGluLeuGlnMetGlnSer 880
Db 2817 GTTCAAAATATGAACACAGAGGAGTATGCTCATAAGATTTTGGAAATTACAGATGCAAGT 2876
QY 881 IleMetGluLysLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAsnPro 900
Db 2877 ATAATGAAAAAGAAAAATGAAAAACCAAGAGAAATATTGCCAAGCATTACAAGATAACCCA 2936
QY 901 SerLeuIleThrPheLeuCysLysAsnCysSerValLeuAlaCysSerGlyGluAspIle 920
Db 2937 TCACATAAATCTTCCTTTTGCAAAACCTGCAGTGTGCTAGCCTGTTCTCTGGGAAGATATC 2996

QY 921 HisValIleGluLysMetHisHisValAsnMetThrProGluPheLysGluLeuTyrIle 940
Db 2997 CATGTAATTGAGAAAAATGCATCAGTCAATATGACCCAGAAATTCAGGAACCTTTACATT 3056
QY 941 ValArgGluAsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluIle 960
Db 3057 GTAAGAGAAAAACAAGCACTGCAAAAGAAGTGTGCCGACTATCAAAATAAATGGTGAATC 3116
QY 961 IleCysLysCysGlyGlnAlaTrpGlyThrMetMetValHisLysGlyLeuAspLeuPro 980
Db 3117 ATCTGCAAAATGTGCCAGGCTTGGGAACAATGATGGTGCACAAAGGCTTAGATTTCCT 3176
QY 981 CysLeuLysIleArgAsnPheValValPheLysAsnAsnSerThrLysLysGlnTyr 1000
Db 3177 TGTCTCAAAATAAGGAATTTTGTAGTGGTTTCAAAAATAATTCACAAAGAAACAATAC 3236
QY 1001 LysLysTrpValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCysLeu 1020
Db 3237 AAAAAGTGGTAGAATTACCTATCACATTTCCCAATCTTGACTATTCAGAATGCTGTTTA 3296
QY 1021 PheSerAspGluAsp 1025
Db 3297 TTTAGTGATGAGGAT 3311
RESULT 5
ADN04879
ID ADN04879 standard; cDNA; 3379 BP.
XX
AC ADN04879;
XX
DT 01-JUL-2004 (first entry)
XX
DE Antipsoriatic cDNA sequence #654.
XX
KW ds; gene; antipsoriatic; gene therapy; psoriasis; diagnosis.
XX
OS Homo sapiens.
XX
PN WO2004028479-A2.
XX
PD 08-APR-2004.
XX
PF 25-SEP-2003; 2003WO-US030907.
XX
PR 25-SEP-2002; 2002US-0414006P.
XX
PA (GETH) GENENTECH INC.
XX
PI Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;
PI Wu TD;
XX
DR WPI; 2004-305105/28.
DR P-PSDB; ADN04880.
XX
PT New PRO nucleic acid or polypeptide, useful for preparing a
PT pharmaceutical composition for diagnosing or treating psoriasis in a
PT mammal.
XX
PS Claim 1; SEQ ID NO 1273; 3069pp; English.
XX
CC The invention relates to novel polynucleotide and polypeptides for
CC treating psoriasis or a sequence having at least 80% identity to the
CC above sequences. The nucleic acid is useful for preparing a composition
CC for diagnosing or treating psoriasis in a mammal. This sequence
CC corresponds to one of the polynucleotides of the invention.
XX
SQ Sequence 3379 BP; 1152 A; 644 C; 753 G; 830 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 0 Length: 3379
Score: 5283.00 Matches: 1024
Percent Similarity: 99.90% Conservative: 0
Best Local Similarity: 99.90% Mismatches: 1

Query Match:	99.47%	Indels:	1
DB:	12	Gaps:	0
US-09-515-363C-2 (1-1025) x ADN04879 (1-3379)			
QY	1	MetSerAsnGlyTyrSerThrAspGluAsnPheArgTyrLeuIleSerCysPheArgAla	20
Db	169	ATGTCGAATGGGTATTCACACAGACGAGAAATTCCTCGCTATCTCATCTCTCGTTCAGGGCC	228
QY	21	ArgValLysMetTyrIleGlnValGluProValLeuAspTyrLeuThrPheLeuProAla	40
Db	229	AGGGTGAATAATGTACATCCAGGTGGAGCCTGTGTGGACTACCTGACCTTTCTGCCTGCA	288
QY	41	GluValLysGluGlnIleGlnArgThrValAlaThrSerGlyAsnMetGlnAlaValGlu	60
Db	289	GAGGTGAAGGAGCAGATTTCAGAGACAGTCGCCACCTCCGGGAACATGCGAGCAGTTGAA	348
QY	61	LeuLeuLeuSerThrLeuGluLysGlyValTrpHisLeuGlyTrpThrArgGluPheVal	80
Db	349	CTGCTGCTGAGCACCTTGGAGAAAGGAGTCTGGCACCTTGGTGGACTCGGGAATTCGTG	408
QY	81	GluAlaLeuArgArgThrGlySerProLeuAlaAlaArgTyrMetAsnProGluLeuThr	100
Db	409	GAGGCCCTCCGAGAACCGGCAGCCCTCTGGCCGCCGCTACATGAACCTGAGCTCAGC	468
QY	101	AspLeuProSerProSerPheGluAsnAlaHisAspGluTyrLeuGlnLeuLeuAsnLeu	120
Db	469	GACTTGGCCCTCTCCATCGTTTGAGAACGCTCATGATGAATATCTCCAACCTGCTGAACCTC	528
QY	121	LeuGlnProThrLeuValAspLysLeuLeuValArgAspValLeuAspLysCysMetGlu	140
Db	529	CTTCAGCCCACTCTGGTGGACAAAGCTTCTAGTAGAGCTCTTGGATAAGTGCATGGAG	588
QY	141	GluGluLeuLeuThrIleGluAspArgAsnArgIleAlaAlaGluAsnAsnGlyAsn	160
Db	589	GAGGAACTGTTGACAAATTGAAGACAGAAACCGGATTGTGCTGCAGAAAAACAATGGAAAT	648
QY	161	GluSerGlyValArgGluLeuLeuLysArgIleValGlnLysGluAsnTrpPheSerAla	180
Db	649	GAATCAGGTGTAAGAGAGAGCTACTAAAAAGGATTGTGCAGAAAGAAAACTGGTTCTCTGCA	708
QY	181	PheLeuAsnValLeuArgGlnThrGlyAsnAsnGluLeuValGlnGluLeuThrGlySer	200
Db	709	TTTCTGAATGTTCTTCGTCAACACAGGAACAATGAACTTGTCCAAGAGTTAACAGGCTCT	768
QY	201	AspCysSerGluSerAsnAlaGluIleGluAsnLeuSerGlnValAspGlyProGlnVal	220
Db	769	GATTGCTCAGAAAAGCAATGCAGAGATTGAGAAATTTATCACAAGTTGATGGTCTCCTCAAGTG	828
QY	221	GluGluGlnLeuLeuSerThrThrValGlnProAsnLeuGluLysGluValTrpGlyMet	240
Db	829	GAAGAGCAACTTCTTCAACCACAGTTTCAGCCAAATCTGGAGAAGGAGTCTGGGSCATG	888
QY	241	GluAsnAsnSerSerGluSerSerPheAlaAspSerSerValValSerGluSerAspThr	260
Db	889	GAGAAATACTCATCAGAATCATCTTTTGCAGATTCTTCTGTAGTTTCAGAAATCAGACACA	948
QY	261	SerLeuAlaGluGlySerValSerCysLeuAspGluSerLeuGlyHisAsnSerAsnMet	280
Db	949	AGTTTGGCAGAAAGGAAGTGTCTAGATGAAAGTCTTGGACATAACAGCAACATG	1008
QY	281	GlySerAspSerGlyThrMetGlySerAspSerAspGluGluAsnValAlaAlaArgAla	300
Db	1009	GGCAGTGATTCAGGCACCATGGGAAGTGATTTCAGATGAAGAGAATGTGGCAGCAAGACA	1068
QY	301	SerProGluProGluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaLeu	320
Db	1069	TCCCCGGAGCCAGAACTCCAGCTCAGGCCTTACCAAATGGAAGTTGCCAGCCAGCCTTG	1128
QY	321	GluGlyLysAsnIleIleCysLeuProThrGlySerGlyLysThrArgValAlaVal	340
Db	1129	GAAGGGAAGAAATATCATCATCTGCCTCCCTACAGGGAGTGGAAAAACCAGAGTGGCTGTT	1188

QY	341	TyrIleAlaLysAspHisLeuAspLysLysLysLysLysAlaSerGluProGlyLysValIle	360
Db	1189	TACATTGCCAAGGATCATTAGACAAGAAGAAAAAGCATCTGAGCCTGGAAAAAGTTATA	1248
QY	361	ValLeuValAsnLysValLeuLeuValGluGlnLeuPheArgLysGluPheGlnProPhe	380
Db	1249	GTTCTTGTCAATAAGGTACTGTCTAGTTGAACAGCTCTTCCGCAAGGAGTTCCCAACCATTT	1308
QY	381	LeuLysLysTrpTyrArgValIleGlyLeuSerGlyAspThrGlnLeuLysIleSerPhe	400
Db	1309	TTGAAGAAATGGTATCGTGTATTATTGGATTAAAGTGGTGATACCCAACTGAAATAATATCATTT	1368
QY	401	ProGluValValLysSerCysAspIleIleSerThrAlaGlnIleLeuGluAsnSer	420
Db	1369	CCAGAAGTTGTCAAGTCTCTGTGATATTATTATCAGTACAGTCAAAATCCTTGAAAACTCC	1428
QY	421	LeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPheSerLeuIle	440
Db	1429	CTCTTAAACTTGGAAAAATGGAGAAGATGCTGGTGTTCATTGTCTCAGACTTTTCCCTCATTT	1488
QY	441	IleIleAspGluCysHisThrAsnLysGluAlaValTyrAsnAsnIleMetArgHis	460
Db	1489	ATCATTGATGAATGTCATCACACCACAA - GAAGCAGTGTATAATAACATCATGAGGCAT	1547
QY	461	TyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysLysGluAsnLysProValIlePro	480
Db	1548	TATTTGATGCAGAAAGTTGAAAAACAATAGACTCAAGAAAGAAAAACAACCCAGTGATTTCCC	1607
QY	481	LeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyAlaThrLysGlnAla	500
Db	1608	CTTCTCTCAGATACTGGGACTAACAGCTTCACCTGGTGTGGAGGGGCCACGAAGCAAGCC	1667
QY	501	LysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThrIleLysThr	520
Db	1668	AAAGCTGAAGAACACATTTTAAAACTATGTGCCAATCTTGATGCATTTACTTTAAAACT	1727
QY	521	ValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluProCysLysLysPheAla	540
Db	1728	GTTAAAGAAAAACCTTGATCAACTGAAAAACCAAAATACAGGAGCCATGCAAGAAGTTTGCC	1787
QY	541	IleAlaAspAlaThrArgGluAspProPheLysGluLysLeuLeuGluIleMetThrArg	560
Db	1788	ATTGCAGATGCAACCAGAGAAGATCCATTTAAAGAGAAAACTTCTAGAAAAATAATGACAAAG	1847
QY	561	IleGlnThrTyrCysGlnMetSerProMetSerAspPheGlyThrGlnProTyrGluGln	580
Db	1848	ATTCAAACTTATTGTCAAATGAGTCCAAATGTCCAGATTTTGGAACTCAACCCCTATGAACAA	1907
QY	581	TrpAlaIleGlnMetGluLysLysAlaAlaLysLysGlyAsnArgLysGluArgValCys	600
Db	1908	TGGGCCATTCAAATGGAAAAAAAAGCTGCAAAAAAAGGAAATCGCAAAAGAACGTGTTGT	1967
QY	601	AlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMetIle	620
Db	1968	GCAGAACATTTGAGGAAGTACAATGAGGCCCTACAAATTAATGACACAAATTCGAATGATA	2027
QY	621	AspAlaTyrThrHisLeuGluThrPheTyrAsnGluGluLysAspLysPheAlaVal	640
Db	2028	GATGGGTATACTCATCTTGAACCTTTCTATAATGAAGAGAAAGATAAGAAAGTTTGCAGTC	2087
QY	641	IleGluAspAspSerAspGluGlyGlyAspAspGluTyrCysAspGlyAspGluAspGlu	660
Db	2088	ATAGAAGATGATAGTGTGAGGGTGGTGTGATGATGAGTATTGTGTGATGGTGAAGATGAG	2147
QY	661	AspAspLeuLysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeuPhe	680
Db	2148	GATGATTTAAAGAAACCTTTGAAACTGGATGAAACAGATAGATTTCTCATGACTTTTATTT	2207
QY	681	PheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGluLys	700
Db	2208	TTTTGAAAAACAATAAAATGTTGAAAAGGCTGGCTGAAAAACCCAGAAATATGAAAAATGAAAG	2267
QY	701	LeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArgThrGluGluSerAlaArg	720

Db 2268 CTGACCAATTAAGAAATACCATATGGAGCAATATATAGGACTGAGGAATCAGCACGA 2327
QY 721 GlyIleIlePheThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTyrIleThrGlu 740
Db 2328 GGAATAATCTTTACAAAACACGACAGAGTGCATATGCGTTTCCAGTGGATTACTGAA 2387
QY 741 AsnGluLysPheAlaGluValGlyValLysAlaHisHisLeuIleGlyAlaGlyHisSer 760
Db 2388 AATGAAAAATTTGCTGAAGTAGGAGTCAAAAGCCCAATCTGATTGGAGCTGGACACAGC 2447
QY 761 SerGluPheLysProMetThrGlnAsnGluGlnLysGluValIleSerLysPheArgThr 780
Db 2448 AGTGAGTTCAAAACCCCATGACACAGATGAACAAAAAGAGTCATTAGTAAATTCGCACT 2507
QY 781 GlyLysIleAsnLeuLeuIleAlaThrThrValAlaGluGlyLeuAspIleLysGlu 800
Db 2508 GGAAAAATCAATCTGCTTATCGCTACACAGTGGCAGAGAAGGTCTGGATATTAAAGAA 2567
QY 801 CysAsnIleValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAlaArg 820
Db 2568 TGTAACATTGTTATCCGTTATGTTTCGTCCCAATGAATAGCCATGGTCCAGGCCCGT 2627
QY 821 GlyArgAlaArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGlyVal 840
Db 2628 GGTCGAGCCAGAGCTGATGAGAGCACCTACGTCCTGGTGTCTCACAGTGGTTACGAGTT 2687
QY 841 IleGluHisGluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHisCys 860
Db 2688 ATCGAACATGAGACAGTTAATGATTCCCGAGAGAGATGATGTATAAAGCTATACATTGT 2747
QY 861 ValGlnAsnMetLysProGluGluTyrAlaHisLysIleLeuGluLeuGlnMetGlnSer 880
Db 2748 GTTCAAAATATGAACACAGAGGAGTATGCTCATAAGATTTTGGAAATTACAGATGCAAGT 2807
QY 881 IleMetGluLysLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAsnPro 900
Db 2808 ATAATGGAAAAAGAAATGAAACCAAGAGAAATATTGCCAGCATTACAAGAAATACCCA 2867
QY 901 SerLeuIleThrPheLeuCysLysAsnCysSerValLeuAlaCysSerGlyGluAspIle 920
Db 2868 TCACATAATACTTCCTTTGGCAAAACTGCAGTGTGCTAGCTGTCTCTGGGGAAGATATC 2927
QY 921 HisValIleGluLysMetHisHisValAsnMetThrProGluPheLysGluLeuTyrIle 940
Db 2928 CATGTAATTGAAAAAATGCATCACGTCAATATGACCCAGAAATCAAGGAACCTTACATT 2987
QY 941 ValArgGluAsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluIle 960
Db 2988 GTAAGAGAAAAACAAAGCACTGCAAAAGAAAGTGTCCGACTATCAAAATAAATGGTGAATC 3047
QY 961 IleCysLysCysGlyGlnAlaTyrGlyThrMetMetValHisLysGlyLeuAspLeuPro 980
Db 3048 ATCTGCAAATGTGGCCAGGCTTGGGGAACAATGATGGTGCACAAAGGCTTAGATTGCTT 3107
QY 981 CysLeuLysIleArgAsnPheValValPheLysAsnAsnSerThrLysLysGlnTyr 1000
Db 3108 TGTCTCAAATAAGGAATTTTGTAGTGGTTTTCAAAAATAATTCAACAAAGAAACAATAC 3167
QY 1001 LysLysTyrValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCysLysLeu 1020
Db 3168 AAAAAGTGGGTAGAATTACCTATCATATTTCCCAATCTTGACTATTTCAGAAATGCTGTTA 3227
QY 1021 PheSerAspGluAsp 1025
Db 3228 TTTAGTGAATGAGGAT 3242
RESULT 6
AAD11170
ID AAD11170 standard; DNA; 3131 BP.
XX
AC AAD11170;
XX

DT 29-NOV-2001 (first entry)
XX Human melanoma differentiation associated-5 protein-related DNA.
DE Human melanoma differentiation associated gene; Mda-5; interferon; IFN;
KW RNA helicase motif; caspase recruitment domain; CARD; therapy; melanoma;
KW neuroblastoma; astrocytoma; glioblastoma; multiforme; cancer; cervical;
KW breast; colon; prostate; osteosarcoma; chondrosarcoma; systemic toxicity;
KW central nervous system; cytostatic; apoptosis; ds.
XX Homo sapiens.
OS
XX WO200164707-A1.
PN
XX 07-SEP-2001.
PD
XX 28-FEB-2001; 2001WO-US006960.
PF
XX 29-FEB-2000; 2000US-00515363.
PR (UYCO) UNIV COLUMBIA NEW YORK.
XX
PI Fisher PB, Kang D, Gopalkrishnan RV;
XX WPI; 2001-565494/63.
DR
XX Nucleic acid sequences encoding a Melanoma Differentiation Associated
PT Gene useful for cancer cell growth suppression, apoptosis and anti-viral
PT activity.
XX
PS Disclosure; Page 134-148; 152pp; English.
XX
CC The present invention relates to an isolated nucleic acid encoding a
CC melanoma differentiation associated gene-5 (Mda-5) polypeptide. Mda-5
CC contains a caspase recruitment domain (CARD) and a RNA helicase motif.
CC Mda-5 is a novel interferon (IFN) inducible gene with structural
CC similarities to RNA helicases and CARD motif containing proteins. Mda-5
CC is induced during terminal differentiation in human melanoma cells
CC treated with the combination of recombinant fibroblast IFN and the
CC antileukaemic compound mezerein (MEZ). Mda-5 is useful for identifying
CC compounds that may induce its expression. Mda-5 is useful for treating
CC cancer e.g., melanoma, neuroblastoma, astrocytoma, glioblastoma
CC multiforme, cervical cancer, breast cancer, colon cancer, prostate
CC cancer, osteosarcoma, chondrosarcoma and especially where the cancer is a
CC cancer of the central nervous system and apoptosis. The Mda-5 promoter
CC exhibits melanocyte tissue specificity and minimises systemic toxicity.
CC The present sequence is human Mda-5 protein-related DNA. Note: The
CC present sequence is designated SEQ ID NO:2 in the sequence listing, but
CC does not correspond to the sequence designated SEQ ID NO:2 in the main
CC body of the specification (AAE10155). The present sequence is not further
CC referred to in the specification, and has been represented in a protein
CC format in the sequence listing
XX
SQ Sequence 3131 BP; 1075 A; 583 C; 699 G; 774 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 0 Length: 3131
Score: 5276.50 Matches: 1024
Percent Similarity: 99.71% Conservative: 0
Best Local Similarity: 99.71% Mismatches: 0
Query Match: 99.35% Indels: 3
DB: 4 Gaps: 1

US-09-515-363C-2 (1-1025) x AAD11170 (1-3131)

QY 1 MetSerAsnGlyTyrSerThrAspGluAsnPheArgTyrLeuIleSerCysPheArgAla 20
Db 1 ATGTGGAATGGGTATTCACACAGACGAGAATTTCCGCTATCTCATCTCGTGTTCAGGGCC 60
QY 21 ArgValLysMetTyrIleGlnValGluProValLeuAspTyrLeuThrPheLeuProAla 40
Db 61 AGGGTGAATAATGTACATCCAGGTGGAGCCTGTGTGGACTACCTGACCTTTCTGCCTGCA 120

Db 2278 GCAGTGAGTTCAAACCCCATGACACAGAAATGAACAAAAAGAGTCATTAGTAAATTCGCA 2337

QY hrGlyLysIleAsnLeuLeuIleAlaThrValAlaGluGlyLeuAspIleLysG 800

Db 2338 CTGGAAAAATCAATCTGCTTATCGCTACACACAGTGGCAGAGAGGTCTGGATATTAAAG 2397

QY luCysAsnIleValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAlaA 820

Db 2398 AATGTAACATTGTTATCCGTTATGGTCTCGTCACCAATGAATAGCCATGGTCCAGGCC 2457

QY rgGlyArgAlaArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGlyV 840

Db 2458 GTGGTCGAGCCAGAGCTGATGAGAGCAGCTACGTCTCTGGTTGCTCACAGTGGTTCAGGAG 2517

QY alIleGluHisGluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHisC 860

Db 2518 TTATCGAACATGAGACAGTTAATGATTCCGAGAGAGATGATGTATAAAGCTATACATT 2577

QY ysValGlnAsnMetLysProGluGluTyrAlaHisLysIleLeuGluLeuGlnMetGlnS 880

Db 2578 GTGTTCAAATATGAACACCAGAGGAGTATGCTCATAGATTTTGGAAATTACAGATGCAAA 2637

QY erIleMetGluLysLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAsnP 900

Db 2638 GTATAATGGAAGAAAAATGAAACCAAGAGAAATATTGCCAAGCATTACAAGATAACC 2697

QY roSerLeuIleThrPheLeuCysLysAsnCysserValLeuAlaCysSerGlyGluAspI 920

Db 2698 CATCACTAATAACTTTCCTTTGCAAAAAGTGCAGTGTGCTAGCTGTCTCTGGGGAAGATA 2757

QY leHisValIleGluLysMetHisHisValAsnMetThrProGluPheLysGluLeuTyrI 940

Db 2758 TCCATGTAATTGAGAAATGCATCAGTCATGATGACCCCAAGAAATTCAGGAACTTTACA 2817

QY leValArgGluAsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluI 960

Db 2818 TTGTAAGAGAAAAACAAAGCAGCTGCAAAAGAAAGTGTGCCGACTATCAATAAATGGTGAAA 2877

QY leIleCysLysCysGlyGlnAlaTrpGlyThrMetMetValHisLysGlyLeuAspLeuP 980

Db 2878 TCATCTGCAAAATGTGGCCAGGCTTGGGGAACAATGATGGTGCAAAAGGCTTAGATTTC 2937

QY roCysLeuLysIleArgAsnPheValValPheLysAsnAsnSerThrLysLysGlnT 1000

Db 2938 CTTGTCTCAAAATAAGGAATTTTGTAGTGGTTTTCAAAAAATAATTCAACAAAGAAACAAT 2997

QY YrLysLysTrpValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCysCysL 1020

Db 2998 ACAAAAAGTGGGTAGAAATTACCTATCACATTTCCCAATCTTGACTATTTCAGAAATGCTGT 3057

QY 1020 euPheSerAspGluAsp 1025

Db 3058 TATTTAGTGATGAGGAT 3074

RESULT 7

ADQ22237

ID ADQ22237 standard; DNA; 3668 BP.

XX AC ADQ22237;

XX 26-AUG-2004 (first entry)

DT Human soft tissue sarcoma-upregulated DNA - SEQ ID 5057.

DE soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human; ds.

XX Homo sapiens.

OS WO2004048938-A2.

XX 10-JUN-2004.

PD

XX

PF 26-NOV-2003; 2003WO-US038193.

XX

PR 26-NOV-2002; 2002US-0429739P.

XX

PA (PROT-) PROTEIN DESIGN LABS INC.

XX

PI Aziz N, Ginsburg WM, Zlotnik A;

XX WPI; 2004-441208/41.

DR

XX

PT Early detection of soft tissue sarcoma comprises determining expression of a gene in a first soft tissue sample and a normal soft tissue sample and comparing the gene expression, also useful in treating soft tissue sarcoma.

PS

XX Example 2; SEQ ID NO 5057; 210pp; English.

XX

CC The invention relates to a novel method for detecting soft tissue sarcoma which comprises obtaining a first soft tissue sample from an individual and a normal soft tissue sample from the same or different individual, determining the expression of a gene in both samples and comparing the expression of the gene in both soft tissue samples, where a higher level of protein expression in the first soft tissue sample indicates the presence of soft tissue sarcoma. The method of the invention has cytostatic applications and may be useful for detecting soft tissue sarcoma, possibly via gene therapy or vaccine production. The nucleic acid sequences may be useful in diagnostic and screening applications. The current sequence is that of a human soft tissue sarcoma-upregulated DNA of the invention. The current sequence is not shown within the specification per se but was submitted in CD format by the inventor.

XX

SQ Sequence 3668 BP; 1215 A; 730 C; 811 G; 872 T; 0 U; 40 Other;

Alignment Scores:

Pred. No.: 0 Length: 3668

Score: 5251.00 Matches: 1015

Percent Similarity: 99.12% Conservative: 1

Best Local Similarity: 99.02% Mismatches: 9

Query Match: 98.87% Indels: 0

DB: 12 Gaps: 0

US-09-515-363C-2 (1-1025) x ADQ22237 (1-3668)

QY 1 MetSerAsnGlyTyrSerThrAspGluAsnPheArgTyrLeuIleSerCysPheArgAla 20

Db 169 ATGTCGAATGGGTATTCCACAGACGAGAATTTCCGCTATCTCATCTCGTGTTCAGGGCC 228

QY 21 ArgValLysMetTyrIleGlnValGluProValLeuAspTyrLeuThrPheLeuProAla 40

Db 229 AGGGTGAAAAATGTACATCCAGGTGGAGCCTGTGTGGACTACCTGACCTTTCTGCCTGCA 288

QY 41 GluValLysGluGlnIleGlnArgThrValAlaThrSerGlyAsnMetGlnAlaValGlu 60

Db 289 GAGGTGAAGGAGCAGATTTCAGAGGACAGTCGCCACCTCCGGGAACATGCAGGCAGTTGAA 348

QY 61 LeuLeuLeuSerThrLeuGluLysGlyValTrpHisLeuGlyTrpThrArgGluPheVal 80

Db 349 CTGCTGCTGAGCACCTTGGAGAAAGGAGTCTGGCACCTTGGTGGACTCGGGAATTCGTG 408

QY 81 GluAlaLeuArgArgThrGlySerProLeuAlaAlaArgTyrMetAsnProGluLeuThr 100

Db 409 GAGGCCCTCCGGAGAACCCGGCAGCCCTCTGGCCGCCCTACATGAACCCCTGAGCTCAG 468

QY 101 AspLeuProSerProSerPheGluAsnAlaHisAspGluTyrLeuGlnLeuLeuAsnLeu 120

Db 469 GACTTGCCCTCTCCATCGTTTGAGAACGCTCATGATGAATATCTCCAACCTGCTGAACCTC 528

QY 121 LeuGlnProThrLeuValAspLysLeuLeuValArgAspValLeuAspLysCysMetGlu 140

Db 529 CTTTCAGCCCACTCTGGTGGACAGCTTCTAGTTAGAGCGTCTTGGATAAGTGCATGGAG 588

QY 141 GluGluLeuLeuThrIleGluAspArgAsnArgIleAlaAlaGluAsnAsnGlyAsn 160

QY 881 IleMetGluLysLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAsnPro 900
Db 2809 ATAATGGAAGAAATGAAACCAAGAGAAATATTGCCAAGCATTAAGAATAACCCA 2868
QY 901 SerLeuIleThrPheLeuCysLysAsnCysSerValLeuAlaCysSerGlyGluAspIle 920
Db 2869 TCACTAATAACTTTCCTTTGCAAAACCTGAGTGTCTAGCCCTGTTCTGGGAAGATATC 2928
QY 921 HisValIleGluLysMetHisHisValAsnMetThrProGluPheLysGluLeuTyrIle 940
Db 2929 CATGTAATTGAGAAATGTCATCAGTCAATATGACCCAGAAATCAAGGAACCTTTACATT 2988
QY 941 ValArgGluAsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluIle 960
Db 2989 GTAAGAGAAAACAAAGCACTGCAAAAGAAAGTGTCCGACTATCAAAATAATGGTGAATC 3048
QY 961 IleCysLysCysGlyGlnAlaTyrGlyThrMetMetValHisLysGlyLeuAspLeuPro 980
Db 3049 ATCTGCAATGTGGCCAGGCTTGGGGAACAATGATGGTGCACAAAGGCTTAGATTGGCT 3108
QY 981 CysLeuLysIleArgAsnPheValValPheLysAsnAsnSerThrLysLysGlnTyr 1000
Db 3109 TGTCTCAAAATAAGGAATTTGTAGTGGTTTCAAAAATAATCAACAAAGAAACATAC 3168
QY 1001 LysLysTrpValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCysCysLeu 1020
Db 3169 AAAAAGTGGTAGAATTACCTATCACATTTCCCAATCTTGACTATTTCAGAAATGCTGTTA 3228
QY 1021 PheSerAspGluAsp 1025
Db 3229 TTTAGTGATGAGGAT 3243

RESULT 8

ADJ75813

ID ADJ75813 standard; DNA; 3771 BP.

XX

AC ADJ75813;

XX

DT 20-MAY-2004 (first entry)

XX

DE Marker gene.SEQ ID NO:1065.

XX

KW bronchial asthma; chronic obstructive pulmonary disease;

KW respiratory epithelial cell; interleukin-13; respiratory; antiasthmatic;

XX gene therapy; marker gene; gene; ds.

OS Mus musculus.

XX

PN EPI394274-A2.

XX

PD 03-MAR-2004.

XX

PF 04-AUG-2003; 2003EP-00254857.

XX

PR 06-AUG-2002; 2002JP-00229312.

PR 20-MAR-2003; 2003JP-00077212.

XX

PA (GENO-) GENOX RES INC.

XX

PI Ohtani N, Sugita Y, Yamaya M, Kubo H, Nagai H, Izuhara K;

XX WPI; 2004-193155/19.

DR

PT Testing for bronchial asthma or chronic obstructive pulmonary disease by

PT comparing the expression level of a marker gene in a biological sample

PT from a subject with the expression level of the gene in a sample from a

PT healthy subject.

XX

PS Claim 14; SEQ ID NO 1065; 241pp; English.

XX

CC The present invention describes a method of testing for bronchial asthma

CC or chronic obstructive pulmonary disease. The method comprises

CC determining the expression level of a marker gene in a biological sample

CC

CC from a subject, comparing the expression level determined with the
CC expression level of the marker gene in a biological sample from a healthy
CC subject, and judging whether the subject has bronchial asthma or chronic
CC obstructive pulmonary disease. The marker gene comprises: (a) a group of
CC genes (S1) whose expression levels increase when respiratory epithelial
CC cells are stimulated with interleukin-13; or (b) a group of genes (S2)
CC whose expression levels decrease when respiratory epithelial cells are
CC stimulated with interleukin-13. Also described: (1) a reagent (I) for
CC testing for bronchial asthma or chronic obstructive pulmonary disease;
CC (2) a kit for screening for a candidate compound for a therapeutic agent
CC to treat bronchial asthma or chronic obstructive pulmonary disease; (3)
CC an animal model for bronchial asthma or chronic obstructive pulmonary
CC disease; (4) an inducer that induces bronchial asthma in a mouse; (5) a
CC method for producing an animal model for bronchial asthma or chronic
CC obstructive pulmonary disease; (6) a therapeutic agent for bronchial
CC asthma or chronic obstructive pulmonary disease, comprising the compound,
CC a marker gene or an antisense nucleic acid corresponding to a portion of
CC the marker gene, a ribozyme, a polynucleotide that suppresses the
CC expression of the gene through an RNAi effect or an antibody recognising
CC a protein encoded by a marker gene; and (7) a DNA chip for testing for
CC bronchial asthma or a chronic obstructive pulmonary disease, on which a
CC probe has been immobilised to assay a marker gene. (I) has respiratory
CC and antiasthmatic activities, and can be used in gene therapy. The method
CC is useful for testing for or screening for a therapeutic agent for
CC bronchial asthma or chronic obstructive pulmonary disease. The present
CC sequence is used in the exemplification of the present invention.

SQ Sequence 3771 BP; 1183 A; 821 C; 866 G; 901 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 0 Length: 3771
Score: 4194.00 Matches: 817
Percent Similarity: 86.94% Conservative: 75
Best Local Similarity: 79.63% Mismatches: 132
Query Match: 78.97% Indels: 2
DB: 12 Gaps: 2

US-09-515-363C-2 (1-1025) x ADJ75813 (1-3771)

QY 1 MetSerAsnGlyTyrSerThrAspGluAsnPheArgTyrLeulleSerCysPheArgAla 20
Db 235 ATGTCGATTGTCTGTTCTGCAGAGGACAGCTTCAGGAATCTCATTTATTTTCAGGCC 294
QY 21 ArgValLysMetTyrIleGlnValGluProValLeuAspTyrLeuThrPheLeuProAla 40
Db 295 AGGCTGAAATGTACATTTCAGTGGAGCCAGTGTGGACACACCTCATCTTTCTGTGCA 354
QY 41 GluValLysGluGlnIleGlnArgThrValAlaThrSerGlyAsnMetGlnAlaValGlu 60
Db 355 GAAACCAAGAGCAGATTCTTAAAGATCAACACCTTGGTAAACACCAGCGGCGAGAA 414
QY 61 LeuLeuLeuSerThrLeuGluLysGlyValTrpHisLeuGlyTrpThrArgGluPheVal 80
Db 415 CTGCTGCTGAGCACCTTGGAGCAGGACAATGGCCCTCTGGGATGGACGAGATGTCGTG 474
QY 81 GluAlaLeuArgArgThrGlySerProLeuAlaAlaArgTyrMetAsnProGluLeuThr 100
Db 475 GAGGCCCTAGAGCACAGTGGCAATCCCTAGCCGCGCTATGTCAAACCCACACTCACT 534
QY 101 AspLeuProSerProSerPheGluAsnAlaHisAspGluTyrLeuGlnLeuLeuAsnLeu 120
Db 535 GATCTGCCCTCTCCTTCTCTGAGACTGCCCATGACGAGTGTCTCCACTTGTGACCCCTC 594
QY 121 LeuGlnProThrLeuValAspLysLeuLeuValArgAspValLeuAspLysCysMetGlu 140
Db 595 CTCCAGCCCACTTGGTGGACAAACTTCTGATTAACGATGTCTTGGACACTTGTTCGAG 654
QY 141 GluGluLeuLeuThrIleGluAspArgAsnArgIleAlaAlaGluAsnAsnGlyAsn 160
Db 655 AAGGACTATTGACAGTCGAGACAGAAATCGGATTCTGTGCTGCAGGAAACAGCGGAAT 714
QY 161 GluSerGlyValArgGluLeuLeuLysArgIleValGlnLysGluAsnTrpPheSerAla 180

Db 715 GAGTCAGGTGTAAGAGAGCTGCTGAGAAAGGATTGTGCAGAAAGGAAAACTGGTTTCTTACC 774
Qy 181 PheLeuAsnValLeuArgGlnThrGlyAsnAsnGluLeuValGlnGluLeuThrGlySer 200
Db 775 TTCTCGGATGTTCTGCGCCAAACTGGAATGATGCATATTCCAAGAATAACAGGTGGA 834
Qy 201 AspCysSerGluSerAsnAlaGluIleGluAsnLeuSerGlnValAspGlyProGlnVal 220
Db 835 GGCTGCCAGAAGACAAACACAGACTTGGCTAACTCGTCTCACAGAGATGGGCTGCAGCT 894
Qy 221 GluGluGlnLeuLeuSerThrThrValGlnProAsnLeuGluLysGluValTrpGlyMet 240
Db 895 AATGAGTGTCTTCTGCCTGCTGCGATGAGTCAAGTGTGGAGACAGAGGCTGGAACGTA 954
Qy 241 GluAsnAsnSerSerGluSerSerPheAlaAspSerSerValValSerGluSerAspThr 260
Db 955 GACGACATATTACCAGAGGCTTCTGTACAGATTCTGTGTGACCACAGAAATCAGACACA 1014
Qy 261 SerLeuAlaGluGlySerValSerCysLeuAspGluSerLeuGlyHisAsnSerAsnMet 280
Db 1015 AGTTTGGCAGAAGGAAGTGTACAGTGTCTTCGATGAAAGTCTTGACATAACAGCAACATG 1074
Qy 281 GlySerAspSerGlyThrMetGlySerAspSerAspGluGluAsnVal---AlaAlaArg 299
Db 1075 GGCAGGATTCAGGCACCATGGGAAGTGATTTCAGATGAAAGTGTCTATCCACACAAAAAGA 1134
Qy 300 AlaSerProGluProGluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAla 319
Db 1135 GTATCCCGGAGCCAGAACTGCAGCTCAGGCCTTACCAAAATGGAAGTGGCCCAACCAGCT 1194
Qy 320 LeuGluGlyLysAsnIleIleCysLeuProThrGlySerGlyLysThrArgValAla 339
Db 1195 CTAGATGGGAAGAATATTATTATCTGCCTCCCCACGGGAGTGGAAAAACCAGAGTGGCT 1254
Qy 340 ValTyrIleAlaLysAspHisLeuAspLysLysLysAlaSerGluProGlyLysVal 359
Db 1255 GTTTACATCACCAAAGATCACTTAGACAAGAAGCAGGCATCTGAATCCGGGAAGGTT 1314
Qy 360 IleValLeuValAsnLysValLeuLeuValGluGlnLeuPheArgLysGluPheGlnPro 379
Db 1315 ATCGTTCTGTCAATAAGGTAATGTTAGCAGAACAACTTTCCGAAAAGAGTTCAACCCA 1374
Qy 380 PheLeuLysLysTrpTyrArgValIleGlyLeuSerGlyAspThrGlnLeuLysIleSer 399
Db 1375 TATTTGAAGAAATGGTATCGAATTATTGGATTAAAGTGGCGATACCCAGCTGAAAATATCA 1434
Qy 400 PheProGluValValLysSerCysAspIleIleIleSerThrAlaGlnIleLeuGluAsn 419
Db 1435 TTTCCAGAAGTTGTCAAATCTTACGATGTTATTATACGACTGCTCAAATCCTTGAAAAAC 1494
Qy 420 SerLeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPheSerLeu 439
Db 1495 TCCCTCTTAAATCTGGAGAGTGGAGACGATGACGGTGTGCAGCTGTTCAGACTTCTCTCTC 1554
Qy 440 IleIleIleAspGluCysHisHisThrAsnLysGluAlaValTyrAsnAsnIleMetArg 459
Db 1555 ATTATCATTTGATGAGTGCCATCACACCAACCAAGGAGCGAGTCTATAACAACATCATGAGA 1614
Qy 460 HistyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysLysGluAsnLysProValIle 479
Db 1615 CGATAITTTGAAGCAGAAGCTGAGAAACAATGACCTCAAGAAACAAACAAACAGCCATT 1674
Qy 480 ProLeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyGlyAlaThrLysGln 499
Db 1675 CCCCTGCCGAGATACTAGGACTGACAGCTTCACCTGGTGTGGAGCAGCCCAAAAGCAG 1734
Qy 500 AlaLysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThrIleLys 519
Db 1735 TCTGAGGCTGAAAAACATATTTTAAATATATGTGCCAATCTTTGATGCTTTTACCATTAAA 1794
Qy 520 ThrValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluProCysLysLysPhe 539
Db 1795 ACAGTGAAGAGAATCTTGGTCAACTCAAAACACCAAATAAAGGAACCATGTCAAGAAATTT 1854

Qy 540 AlaIleAlaAspAlaThrArgGluAspPropheLysGluLysLeuLeuGluIleMetThr 559
Db 1855 GTGATTGCTGATGACACCAGAGAAAATCCATTTAAAGAGAAAACCTTCTAGAAAATTATGGCA 1914
Qy 560 ArgIleGlnThrTyrCysGlnMetSerProMetSerAspPheGlyThrGlnProTyrGlu 579
Db 1915 AGCATTCAGACTTACTGCCAAAAAAGTCCAATGTTCAGATTTTGGAAACCCCAACATTATGAG 1974
Qy 580 GlnTrpAlaIleGlnMetGluLysLysAlaAlaLysLysGlyAsnArgLysGluArgVal 599
Db 1975 CAGTGGCCATTCAAAATGGAGAAAAAAGCTGTAAAGACGGAATTCGCAAGATCGCGTC 2034
Qy 600 CysAlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMet 619
Db 2035 TGTGCAGAACATTTGAGGAAGTACAACGAAAGCCCTACAAATCAACGACACGATCCGAATG 2094
Qy 620 IleAspAlaTyrThrHisLeuGluThrPheTyrAsnGluGluLysAspLysLysPheAla 639
Db 2095 ATTGATGCATATAGCCACCTGGAGGCATTCTACACTGATGAGAAAGAAAGATTTCGCA 2154
Qy 640 ValIleGluAspAspSerAspGluGlyGlyAspAspGluTyrCysAspGlyAspGluAsp 659
Db 2155 GTCCTC--AATGACAGCGACAAGAGTGTGATGACGAGGCCAGCATGTGCAATGACCAACTT 2211
Qy 660 GluAspAspLeuLysLysProLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGlu 679
Db 2212 AAGGCGATGTAAGAAAAATCTTTGAAACTGGACGAAACGGATGAATTTCTCATGAATTTG 2271
Qy 680 PhePheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGlu 699
Db 2272 TTCTTTGATAACAAGAAAAATGTTGAAAAAACTAGCTGAAAAACCCAAAATACGAGAAATGAA 2331
Qy 700 LysLeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArgThrGluGluSerAla 719
Db 2332 AAACCTCATTAATTAAGAAACACGATACCTGGAACAATTACAAAGGTCTGAGGAGTCTCTCC 2391
Qy 720 ArgGlyIleIlePheThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTrpIleThr 739
Db 2392 CGAGGAATTTATTTTACAAAAACACGACAGAGCACCTACGCACCTTCCAGTGGATCATG 2451
Qy 740 GluAsnGluLysPheAlaGluValGlyValLysAlaHisHisLeuIleGlyAlaGlyHis 759
Db 2452 GAAAAATGCAAAGTTTTCGGGAAGTTGGAGTCAAAGCGCATCACCTGATTTGGCGGGGCAC 2511
Qy 760 SerSerGluPheLysProMetThrGlnAsnGluGlnLysGluValIleSerLysPheArg 779
Db 2512 AGCAGTGAAGTCAAGCCCATGACTCAGACTGAAACAAAAAAGAGTCAATAGTAAATTTTCG 2571
Qy 780 ThrGlyLysIleAsnLeuLeuIleAlaThrThrValAlaGluGluGlyLeuAspIleLys 799
Db 2572 ACTGGCGAAATAAATCTGCTTATCGCTACGACGGTGGCAGAGGAGGCTGGATATCAAA 2631
Qy 800 GluCysAsnIleValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAla 819
Db 2632 GAGTGCAATATTGTTATTGTTATGTCCTTATGGCCCTGTACGAACGAGATAGCCATGGTCCAGGCC 2691
Qy 820 ArgGlyArgAlaArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGly 839
Db 2692 CGGGGTTCGAGCCAGAGCTGATGAAAGCACCTGATGTCTCTGTCACCCAGAGTGGCTCAGGA 2751
Qy 840 ValIleGluHisGluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHis 859
Db 2752 GTTACCGAAACGGGAGATTGTTAATGATTTCGAGAGAAGATGATGTATAAAGCTATTAAAC 2811
Qy 860 CysValGlnAsnMetLysProGluGluTyrAlaHisLysIleLeuGluLeuGlnMetGln 879
Db 2812 CGTGTTCAAAAACATGAAACCAGAGGAGTATGCACATAAGATTTTGGAAATTCAGGTGCAA 2871
Qy 880 SerIleMetGluLysLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAsn 899
Db 2872 AGTATCTTGAAAAAGAAAAATGAAAGTCAAAGAAAGAGCATTTGCAAAAGCAATACACGACAAT 2931

QY	900	ProSerLeuIleThrPheLeuCysLysAsnCysSerValLeuAlaCysSerGlyGluAsp	919	PR	14-AUG-2000;	2000US-0225213P.
Db	2932	CCATCGTTAATAACACTTCTCTGCAAAATTTGTAGCATGCTGGTCTCGGGAGAAAC	2991	PR	14-AUG-2000;	2000US-0225214P.
QY	920	IleHisValIleGluLysMetHisHisValAsnMetThrProGluPheLysGluLeuTyr	939	PR	14-AUG-2000;	2000US-0225266P.
Db	2992	ATCCATGTCATTGAGAAGATGCATCATGTCAATATGACACCAGAATTC AAGGGACTCTAC	3051	PR	14-AUG-2000;	2000US-0225267P.
QY	940	IleValArgGluAsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGlu	959	PR	14-AUG-2000;	2000US-0225268P.
Db	3052	ATTGTAAGAGAAAACAAAGCACTGCAAAAGAAATTTGCTGATTATCAGACCAATGGAGAG	3111	PR	14-AUG-2000;	2000US-0225270P.
QY	960	IleIleCysLysCysGlyGlnAlaTirpGlyThrMetMetValHisLysGlyLeuAspLeu	979	PR	14-AUG-2000;	2000US-0225447P.
Db	3112	ATTATCTGCAAGTGTGCCAGGCTTGGGGAACAATGATGGTGCACAAAGGTTTAGATTG	3171	PR	14-AUG-2000;	2000US-0225757P.
QY	980	ProCysLeuLysIleArgAsnPheValValPheLysAsnAsnSerThrLysLysGln	999	PR	14-AUG-2000;	2000US-0225758P.
Db	3172	CCTTGCTTAAATAAGGAATTTTGTAGTCAATTTCAAAAATAACTACCGAAGAAACAG	3231	PR	14-AUG-2000;	2000US-0225759P.
QY	1000	TyrLysLysTrpValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCysCys	1019	PR	18-AUG-2000;	2000US-0226279P.
Db	3232	TACAAGAAGTGGTGAATTCCTATCAGATTCCTGATCTTGACTACTCAGAATACTGC	3291	PR	22-AUG-2000;	2000US-0226681P.
QY	1020	LeuPheSerAspGluAsp	1025	PR	22-AUG-2000;	2000US-0226868P.
Db	3292	TTGTATAGTGAAGAT	3309	PR	22-AUG-2000;	2000US-0227182P.
RESULT 9				PR	23-AUG-2000;	2000US-0227009P.
AAS40960				PR	30-AUG-2000;	2000US-0228924P.
ID	AAS40960	standard; cDNA; 1967 BP.		PR	01-SEP-2000;	2000US-0229287P.
XX	AAS40960;			PR	01-SEP-2000;	2000US-0229343P.
AC				PR	01-SEP-2000;	2000US-0229344P.
XX	17-DEC-2001	(first entry)		PR	01-SEP-2000;	2000US-0229345P.
XX	cDNA encoding novel human enzyme polypeptide #176.			PR	05-SEP-2000;	2000US-0229509P.
KW	Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;			PR	06-SEP-2000;	2000US-0230437P.
KW	ligase; hyperproliferative disorder; immunodeficiency disorder;			PR	06-SEP-2000;	2000US-0230438P.
KW	autoimmune disorder; neurological disorder; metabolic disorder;			PR	08-SEP-2000;	2000US-0231242P.
KW	inflammatory disorder; cardiovascular disorder; reproductive disorder;			PR	08-SEP-2000;	2000US-0231243P.
KW	blood-related disorder; infectious disorder; gene therapy; cytostatic;			PR	08-SEP-2000;	2000US-0231244P.
XX	anti arthritic; nephrotropic; anticoagulant; ss.			PR	08-SEP-2000;	2000US-0231413P.
OS	Homo sapiens.			PR	08-SEP-2000;	2000US-0231414P.
XX	WO200155301-A2.			PR	08-SEP-2000;	2000US-0232080P.
PN	02-AUG-2001.			PR	08-SEP-2000;	2000US-0232081P.
PD	17-JAN-2001; 2001WO-US001239.			PR	12-SEP-2000;	2000US-0231968P.
XX	31-JAN-2000; 2000US-0179065P.			PR	14-SEP-2000;	2000US-0232397P.
PR	04-FEB-2000; 2000US-0180628P.			PR	14-SEP-2000;	2000US-0232398P.
PR	24-FEB-2000; 2000US-0184664P.			PR	14-SEP-2000;	2000US-0232399P.
PR	02-MAR-2000; 2000US-0186350P.			PR	14-SEP-2000;	2000US-0232400P.
PR	16-MAR-2000; 2000US-0189874P.			PR	14-SEP-2000;	2000US-0232401P.
PR	17-MAR-2000; 2000US-0190076P.			PR	14-SEP-2000;	2000US-0233063P.
PR	18-APR-2000; 2000US-0198123P.			PR	14-SEP-2000;	2000US-0233064P.
PR	19-MAY-2000; 2000US-0205515P.			PR	14-SEP-2000;	2000US-0233065P.
PR	07-JUN-2000; 2000US-0209467P.			PR	21-SEP-2000;	2000US-0234223P.
PR	28-JUN-2000; 2000US-0214886P.			PR	21-SEP-2000;	2000US-0234274P.
PR	30-JUN-2000; 2000US-0215135P.			PR	25-SEP-2000;	2000US-0234997P.
PR	07-JUL-2000; 2000US-0216647P.			PR	25-SEP-2000;	2000US-0234998P.
PR	07-JUL-2000; 2000US-0216880P.			PR	26-SEP-2000;	2000US-0235484P.
PR	11-JUL-2000; 2000US-0217487P.			PR	27-SEP-2000;	2000US-0235834P.
PR	11-JUL-2000; 2000US-0217496P.			PR	27-SEP-2000;	2000US-0235836P.
PR	14-JUL-2000; 2000US-0218290P.			PR	29-SEP-2000;	2000US-0236327P.
PR	26-JUL-2000; 2000US-0220963P.			PR	29-SEP-2000;	2000US-0236367P.
PR	26-JUL-2000; 2000US-0220964P.			PR	29-SEP-2000;	2000US-0236368P.
PR	14-AUG-2000; 2000US-0224518P.			PR	29-SEP-2000;	2000US-0236369P.
PR	14-AUG-2000; 2000US-0224519P.			PR	29-SEP-2000;	2000US-0236370P.
				PR	02-OCT-2000;	2000US-0236802P.
				PR	02-OCT-2000;	2000US-0237037P.
				PR	02-OCT-2000;	2000US-0237038P.
				PR	02-OCT-2000;	2000US-0237039P.
				PR	02-OCT-2000;	2000US-0237040P.
				PR	13-OCT-2000;	2000US-0239935P.
				PR	13-OCT-2000;	2000US-0239937P.
				PR	20-OCT-2000;	2000US-0240960P.
				PR	20-OCT-2000;	2000US-0241221P.
				PR	20-OCT-2000;	2000US-0241785P.
				PR	20-OCT-2000;	2000US-0241786P.
				PR	20-OCT-2000;	2000US-0241787P.
				PR	20-OCT-2000;	2000US-0241808P.
				PR	20-OCT-2000;	2000US-0241809P.
				PR	20-OCT-2000;	2000US-0241826P.
				PR	01-NOV-2000;	2000US-0244617P.
				PR	08-NOV-2000;	2000US-0246474P.
				PR	08-NOV-2000;	2000US-0246475P.
				PR	08-NOV-2000;	2000US-0246476P.
				PR	08-NOV-2000;	2000US-0246477P.
				PR	08-NOV-2000;	2000US-0246478P.

PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.

(HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Barash SC, Ruben SM;

XX WPI; 2001-465566/50.
XX P-PSDB; AAU23090.

PT Novel polypeptides and polynucleotides useful for diagnosing, preventing,
PT treating neural, immune system, muscular, reproductive, pulmonary,
PT cardiovascular, renal, proliferative disorders and cancerous diseases.

XX Claim 4; SEQ ID NO 186; 1180pp; English.

PS The present invention relates to the isolation of novel human enzyme
XX polypeptides (AAU22915-AAU23814), and the cDNA and genomic sequences
CC encoding them. The enzyme polypeptides of the invention may comprise the
CC functional classes of oxidoreductases, transferases, hydrolases, lyases,
CC isomerases or ligases. The sequences of the invention are useful in the
CC diagnosis, treatment, prevention and/or prognosis of a wide range of
CC disorders including hyperproliferative disorders (e.g. cancer),
CC immunodeficiency disorders (e.g. AIDS) autoimmune disorders (e.g.
CC arthritis), neurological disorders (e.g. Alzheimer's disease), metabolic
CC disorders (e.g. phenylketonuria), inflammatory disorders (e.g. asthma),
CC cardiovascular disorders (e.g. atherosclerosis), blood-related disorders
CC (e.g. haemophilia), reproductive disorders (e.g. infertility) and
CC infectious disorders (e.g. influenza). The polynucleotides of the
CC invention can also be used in gene therapy. AAS40785-AAS41684 represent
CC cDNA sequences encoding for the novel human enzyme polypeptides of the
CC invention. Note: The sequence data for this patent did not form part of

CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1967 BP; 723 A; 344 C; 416 G; 476 T; 0 U; 8 Other;

Alignment Scores:
Pred. No.: 4.69e-251 Length: 1967
Score: 3176.00 Matches: 636
Percent Similarity: 97.11% Conservative: 2
Best Local Similarity: 96.80% Mismatches: 18
Query Match: 59.80% Indels: 5
DB: 4 Gaps: 1

US-09-515-363C-2 (1-1025) x AAS40960 (1-1967)

Qy 294 GluAsnValAlaAlaArgAlaSerProGluProGluLeuGlnLeuArgProTyrGlnMet 313
Db 3 GAGAAATGTGGCAGCAAGAGCATCCCGGAGCCAGAACTCCAGCTCAGGCCTTACCAATG 62
Qy 314 GluValAlaGlnProAlaLeuGluGlyLysAsnIleIleCysLeuProThrGlySer 333
Db 63 GAAGTTGCCCGCAGCCAGCCTTGGAAAGGGAAGAATATCATCATCTGCCTCCCTACAGGAGT 122
Qy 334 GlyLysThrArgValAlaValTyrIleAlaLysAspHisLeuAspLysLysLysAla 353
Db 123 GGAAAAACCAGAGTGGCTGTTTACATTGCCAAGGATCATTAGACAAGAAAAAGCA 182
Qy 354 SerGluProGlyLysValIleValLeuValAsnLysValLeuValGluGlnLeuPhe 373
Db 183 TCTGAGCCTGGAAAAAGTTATAGTTCTTGTCAATAAGGTACTGCTAGTTGAACAGCTCTTC 242
Qy 374 ArgLysGluPheGlnProPheLeuLysLysTrpTyrArgValIleGlyLeuSerGlyAsp 393
Db 243 CGCAAGGAGTTCACCACTTTTGAAGAAATGGTATCGTGTATTGGATTAAAGTGGTGAT 302
Qy 394 ThrGlnLeuLysIleSerPheProGluValValLysSerCysAspIleIleIleSerThr 413
Db 303 ACCCAACTGAAAAATATCATTTCCAGAGTTGTCAAGTCTGTGATATTATTATCAGTACA 362
Qy 414 AlaGlnIleLeuGluAsnSerLeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGln 433
Db 363 GCTCAAAATCCTTGAAAAACTCCCTCTTAAACTTGGAAATGGAGAAAGATGCTGGTGTCAA 422
Qy 434 LeuSerAspPheSerLeuIleIleIleAspGluCysHisThrAsnLysGluAlaVal 453
Db 423 TTGTCAGACTTTTCCCTCATTTATCATTTGATGAATGTCATCACACCAACAAGACGAGTG 482
Qy 454 TyrAsnAsnIleMetArgHisTyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysLys 473
Db 483 TATAATAACATCATGAGGCATTATTGATGAGAAAGTTGAAAGAACTATGTGCAATCTT 542
Qy 474 GluAsnLysProValIleProLeuProGlnIleLeuGlyLeuThrAlaSerProGlyVal 493
Db 543 GAAACAAACCAGTGTATCCCTTCCCTCAGATACTGGGACTAACAGCTTACCTGGTGT 602
Qy 494 GlyGlyAlaThrLysGlnAlaLysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeu 513
Db 603 GGAGGGGCCACGAAGCAAGCAAGCTGAAGAACACACATTTTAAAACTATGTGCCAATCTT 662
Qy 514 AspAlaPheThrIleLysThrValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGln 533
Db 663 GATGCATTACTATTAAACTGTTAAAGAAACCTTGATCAACTGAAACCAACCAATACAG 722
Qy 534 GluProCysLysLysPheAlaIleAlaAspAlaThrArgGluAspProPheLysGluLys 553
Db 723 GAGCATGCAAGAAGTTTGCCATTGCAGATGCACAGAGAGATCCATTAAAGAGAAA 782
Qy 554 LeuLeuGluIleMetThrArgIleGlnThrTyrCysGlnMetSerProMetSerAspPhe 573
Db 783 CTTCTAGAAATAATGACAAGGATTCAAACTTATTGTCAAATGAGTCCCAATGTCAGATT 842
Qy 574 GlyThrGlnProTyrGluGlnTyrAlaIleGlnMetGluLysLysAlaAlaLysLysGly 593

Db 843 GGAACCAACCCCTATGAACAATGGGCCATTCAAATGGAAAAAAGCTGCAAAAGAGGA 902

QY AsnArgLysGluArgValCysAlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIle 613

Db 903 AATCGCAAGAACGCTTTGTGCAGAACATTTGAGGAAGTACAAATGAGGCCCTACAAATT 962

QY AsnAspThrIleArgMetIleAspAlaTyrThrHisLeuGluThrPheTyrAsnGluGlu 633

Db 963 AATGACACAATTTCGAATGATAGATGCGTATACTCATCTTGAAACTTCTATAATGAAGAG 1022

QY LysAspLysLysPheAlaValIleGluAspAspSerAspGluGlyGlyAspAspGluTyr 653

Db 1023 AAAGATAAGAAGTTTGCAGTCATAGAAGATGATAGTATGAGGCTGGTGATGATGAT 1082

QY CysAspGlyAspGluAspGluAspAspLeuLysLysProLeuLysLeuAspGluThrAsp 673

Db 1083 TGTGATGGTGTAGATGAGGATGATTAAAGAAACCTTTGAAACTGGATGAAACAGAT 1142

QY ArgPheLeuMetThrLeuPhePheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsn 693

Db 1143 AGATTTCTCATGACTTTATTTTGGAAAAACAATAAAATGTTGAAAGGCTGGCTGAAAC 1202

QY ProGluTyrGluAsnGluLysLeuThrLysLeuArgAsnThrIleMetGluGlnTyrThr 713

Db 1203 CCAGAATATGAAATGAAAAGCTGACCAATTAAGAAATACCAATGGAGCAATATACT 1262

QY ArgThrGluGluSerAlaArgGlyIleIlePheThrLysThrArgGlnSerAlaTyrAla 733

Db 1263 AGGACTGAGGAATCAGCACGAGGAATAATCTTTACAAAAACACGACAGAGTGCTATGCG 1322

QY LeuSerGlnTrpIleThrGluAsnGluLysPheAlaGluValGlyValLysAlaHisHis 753

Db 1323 CTTTCCCAGTGGATTACTGAAAATGAAAATTTGCTGAAGTAGGAGTCAAAAGCCCCCAT 1382

QY LeuIleGlyAlaGlyHisSerSerGluPheLysProMetThrGlnAsnGluGlnLysGlu 773

Db 1383 CTGATTGGAGCTGGACACAGCAGTGAGTTCAAAACCCATGACACAGAAATGAACAAAGAA 1442

QY ValIleSerLysPheArgThrGlyLysIleAsnLeuLeuIleAlaThrThrValAlaGlu 793

Db 1443 GTCATTAGTAAATTTTCGCACTGGAAAAATAAATCTGCTTATCGTCC-ACAGTGGCAGAA 1501

QY GluGlyLeuAspIleLysGluCysAsnIleValIleArgTyrGlyLeuValThrAsnGlu 813

Db 1502 GAAGGTCTGGATATTAAAGAATGTAACATTGKTATCCGTTATGCTTCGTCACCAATGAA 1561

QY IleAlaMetValGlnAlaArgGlyArgAlaArgAlaAspGluSerThrTyrValLeuVal 833

Db 1562 ATACC-ATGGTCCAGCC-CGTGGTCGAGTAGAGCTGATGAGAGCACCTACGT-CTGGTT 1618

QY AlaHisSerGlySerGlyValIleGluHisGluThrValAsnAspPheArgGluLysMet 853

Db 1619 GCTCACAGTGGTTCAGGAGTTATCGAACGTGAGACAGTTAATGATTTCCGAGAGAGATG 1678

QY MetTyrLysAlaIleHisCysValGlnAsnMetLysProGluGluTyrAlaHisLysIle 873

Db 1679 ATGTATAAGCTATC---ATTGKGTTCAAATATGAACACAGAGGAGTATGCTYATAAGAWT 1735

QY LeuGluLeuGlnMetGlnSerIleMetGluLysLysMetLysThrLysArgAsnIleAla 893

Db 1736 TTGGAATTACAGATGCMWAAGTATAATGGAAGAAAGAAATGAAACCAAGAGAAATATTGCC 1795

QY LysHisTyrLysAsnAsnProSerLeuIleThrPheLeuCysLysAsnCysSerValLeu 913

Db 1796 AAGCATTACAAGATAACCCATCACTAATAACTTTCCTTTGCAAAAACTGCAGTGTGCTA 1855

QY AlaCysSerGlyGluAspIleHisValIleGluLysMetHisHisValAsnMetThrPro 933

Db 1856 GCCTGTTCTGGGGAAGATATCCATGNAAATGANAATAATGCATNACGTCAATATGACCCCA 1915

QY GluPheLysGluLeuTyrIleValArgGluAsnLysAlaLeuGlnLysLys 950

Db 1916 GAATTCAGGGAACCTTTACATTGGAAGAGAGAAAAACAAAGCACTGCAAAAAAGAAA 1966

RESULT 10

ADA53260

ID ADA53260 standard; cDNA; 1776 BP.

XX

AC ADA53260;

XX

DT 20-NOV-2003 (first entry)

XX

DE Human coding sequence, SEQ ID 828.

XX

KW Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;

KW Gene Therapy; human; secretory protein; membrane proteins; cancer;

KW inflammatory disease; osteoporosis; neurological disease; gene; ss.

XX

OS Homo sapiens.

XX

PN EPI293569-A2.

XX

PD 19-MAR-2003.

XX

PF 21-MAR-2002; 2002EP-00006586.

XX

PR 14-SEP-2001; 2001JP-00328381.

PR 24-JAN-2002; 2002US-0350435P.

XX

(HELI-) HELIX RES INST.

PA (REAS-) RES ASSOC BIOTECHNOLOGY.

XX

PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;

PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;

PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;

XX

DR WPI; 2003-395539/38.

DR P-PSDB; ADA54899.

XX

PT New polynucleotides encoding full-length polypeptides, e.g. secretory

PT and/or membrane proteins, useful for developing medicines for diseases in

PT which the gene is involved, or as target molecules for gene therapy.

XX

PS Claim 1; SEQ ID NO 828; 205pp; English.

XX

CC The present invention relates to novel human secretory or membrane

CC proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-

CC ADA54071). The coding sequences are useful in the gene therapy of

CC diseases caused by abnormalities of the proteins, e.g. cancer,

CC inflammatory diseases, osteoporosis or neurological disease.

XX

SQ Sequence 1776 BP; 666 A; 291 C; 363 G; 456 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2.71e-226 Length: 1776

Score: 2874.00 Matches: 546

Percent Similarity: 99.64% Conservative: 3

Best Local Similarity: 99.09% Mismatches: 2

Query Match: 54.11% Indels: 0

DB: 10 Gaps: 0

US-09-515-363C-2 (1-1025) x ADA53260 (1-1776)

QY 475 AsnLysProValIleProLeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGly 494

Db 1 CACAAACCAGTGATTCCCTTCCTCAGATACTGGGACTAACAGCTTACCTGGTGTGGA 60

QY 495 GlyAlaThrLysGlnAlaLysAlaGluHisIleLeuLysLeuCysAlaAsnLeuAsp 514

Db 61 GGGGCCACGAAGCAAGCCAAAGCTGAAGAACACATTTTAAACTATGTGCCAATCTTGAT 120

QY 515 AlaPheThrIleLysThrValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGlu 534

Db 121 GCATTTACTATTAATACTGTTAAAGAAAAACCTTGATCACTGAAAAAACCAATACAGGAG 180

QY 535 ProCysLysLysPheAlaIleAlaAspAlaThrArgGluAspProPheLysGluLysLeu 554

Db 181 CCATGCAAGAGTTTGCCATTTCAGATGCAACCCAGAGAAGATCCATTTAAGAGAACTT 240
QY 555 LeuGluIleMetThrArgIleGlnThrTyrCysGlnMetSerProMetSerAspPheGly 574
Db 241 CTAGAAATAATGACAAGGATTCAAACTTATTGTCAAATGAGTCCAATGTCAGATTTTGA 300
QY 575 ThrGlnProTyrGluGlnTrpAlaIleGlnMetGluLysLysAlaAlaLysLysGlyAsn 594
Db 301 ACTCAACCTATGAACAATGGGCCATTCAAATGGAAAAAAAGCTGCAAAAGAGAAAT 360
QY 595 ArgLysGluArgValCysAlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsn 614
Db 361 CGCAAGAAGCGTGTGTGTCAGAACATTTGAGGAAGTACAATGAGGCCCTACAAATTAAT 420
QY 615 AspThrIleArgMetIleAspAlaTyrThrHisLeuGluThrPheTyrAsnGluGluLys 634
Db 421 GACACAATTGCAATGATAGATGCGTATACTCATCTTGAAACTTTCTATAATGAAGAGAAA 480
QY 635 AspLysLysPheAlaValIleGluAspAspSerAspGluGlyGlyAspAspGluTyrCys 654
Db 481 GATAAGAAGTTGCAGTCATAGAAGATGATGATGAGGTTGGTGATGATGAGTATTGT 540
QY 655 AspGlyAspGluAspGluAspLeuLysLysProLeuLysLeuAspGluThrAspArg 674
Db 541 GATGGTGATGAAGATGAGGATGATTTAAAGAAACCTTTGAAACTGGATGAACAGATAGA 600
QY 675 PheLeuMetThrLeuPhePheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnPro 694
Db 601 TTTCTCATGACTTTATTTTTTGAAAAACAATAAAATGTTGAAAAGCTGGCTGAAAAACCCA 660
QY 695 GluTyrGluAsnGluLysLeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArg 714
Db 661 GAATATGAAATGAAAAGCTGACCAAAATTAAGAAATACCAATATGGAGCAATATACTAGG 720
QY 715 ThrGluGluSerAlaArgGlyIleIlePheThrLysThrArgGlnSerAlaTyrAlaLeu 734
Db 721 ACTGAGGAATCAGCACGAGGAATAATCTTTACAAAAACACGACAGAGTGATGCGCTT 780
QY 735 SerGlnTrpIleThrGluAsnGluLysPheAlaGluValGlyValLysAlaHisLeu 754
Db 781 TCCCAGTGGATTACTGAAAATGAAAAATTTGCTGAAGTAGAGTCAAAGCCCAACCATCTG 840
QY 755 IleGlyAlaGlyHisSerSerGluPheLysProMetThrGlnAsnGluGlnLysGluVal 774
Db 841 ATTGGAGCTGGACACAGCAGTGAGTTCAAACCCCATGACACAGAATGAACAAAAAGAGTC 900
QY 775 IleSerLysPheArgThrGlyLysIleAsnLeuIleAlaThrThrValAlaGluGlu 794
Db 901 ATTAGTAAATTTCCGACTGGAAGAATAAATCTGCTTATCGCTACCACAGTGGCAGAGAA 960
QY 795 GlyLeuAspIleLysGluCysAsnIleValIleArgTyrGlyLeuValThrAsnGluIle 814
Db 961 GGTCTGGATATTAAAGAATGTAAACATTGTTATCCGTTATGCTTCGTCAACCAATGAATA 1020
QY 815 AlaMetValGlnAlaArgGlyArgAlaArgAlaAspGluSerThrTyrValLeuValAla 834
Db 1021 GCCATGGTCCAGGCCCGTGGTCGAGCCAGAGCTGATGAGAGCACCTACGTCCTGGTTGCT 1080
QY 835 HisSerGlySerGlyValIleGluHisGluThrValAsnAspPheArgGluLysMetMet 854
Db 1081 CACAGTGGTTCCAGGAGTTATCGAACCGTGAGACAGTTAATGATTTCGAGAGAGATGATG 1140
QY 855 TyrLysAlaIleHisCysValGlnAsnMetLysProGluGluTyrAlaHisLysIleLeu 874
Db 1141 TATAAAGCTATACATTGTGTTCAAAATATGAAACCCAGAGGAGTATGCTCATAGATTGTG 1200
QY 875 GluLeuGlnMetGlnSerIleMetGluLysLysMetLysThrLysArgAsnIleAlaLys 894
Db 1201 GAATTACAGATGCAAAAGTATAATGGAAGAAAAATGAAACCAAGAGAAATATTGCCAAG 1260
QY 895 HisTyrLysAsnAsnProSerLeuIleThrPheLeuCysLysAsnCysSerValLeuAla 914

Db 1261 CATTACAAGATAACCCATCACTAATAAATTTCTTTTGCAAAAACTGCAGTGTGTAGCC 1320
QY 915 CysSerGlyGluAspIleHisValIleGluLysMetHisValAsnMetThrProGlu 934
Db 1321 TGTTCCTGGGAAGATATCCATGTAATTGAGAAAATGCATCACGTCAATATGACCCAGAA 1380
QY 935 PheLysGluLeuTyrIleValArgGluAsnLysAlaLeuGlnLysLysCysAlaAspTyr 954
Db 1381 TTCAAGGAACTTTACATTGTAAGAGAAAAACAACACTGCAAAAGAAAGTGTGCCGACTAT 1440
QY 955 GlnIleAsnGlyGluIleIleCysLysCysGlyGlnAlaTrpGlyThrMetMetValHis 974
Db 1441 CAAATAAATGGTGAATCATCTGCAATGTGCCAGGCTTGGCGAAACAATGATGTGCAC 1500
QY 975 LysGlyLeuAspLeuProCysLeuLysIleArgAsnPheValValPheLysAsnAsn 994
Db 1501 AAAGGCTTAGATTGCCTTGTCTCAAAATAAGGAATTTTGTAGTGGTTTTCAAAAAATAAT 1560
QY 995 SerThrLysLysGlnTyrLysLysTrpValGluLeuProIleThrPheProAsnLeuAsp 1014
Db 1561 TCAACAAAGAAACAATACAAAAGTGGTAGAATTACCTATCACATTTCCCAATCTTGAC 1620
QY 1015 TyrSerGluCysCysLeuPheSerAspGluAsp 1025
Db 1621 TATTCAGAAATGCTGTTTATTATTAGTGATGAGGAT 1653
RESULT 11
ABA04916
ID ABA04916 standard; DNA; 1443 BP.
XX
AC ABA04916;
XX
DT 05-MAR-2002 (first entry)
XX
DE Human RNA helicase RH16 related DNA sequence #3.
XX
KW Human; RH16; RNA helicase; cytostatic; virucide; anti-HIV;
KW immunosuppressive; immunostimulatory; antirheumatic; antiarthritic;
KW antiarteriosclerotic; osteopathic; antidiabetic; hepatotropic;
KW antiinflammatory; cancer; infection; HIV; hepatitis; genetic disease;
KW autoimmune disease; graft rejection; vaccine; ds.
XX
OS Homo sapiens.
XX
PN WO200185955-A1.
XX
PD 15-NOV-2001.
XX
PF 11-MAY-2001; 2001WO-FR001441.
XX
PR 11-MAY-2000; 2000FR-00006030.
XX
PA (ISTA-) ISTAC.
PA (INSP) INST PASTEUR LILLE.
XX
PI Bahr G, Cocude C, Capron A;
XX
DR WPI; 2002-082898/11.
XX
PT New polypeptide, useful for treating and diagnosing cancer or
PT inflammation, and drug screening, comprises a human polynucleotide
PT homologous to RNA helicase.
XX
PS Example 2; Page 94-95; 114pp; French.
XX
CC The present invention relates to human RH16 (see AAM47798). RH16 is a
CC 116kDa protein and has homology to RNA helicases (DEXH box). RH16 and
CC its coding sequence are useful for treating cancer; acute or chronic
CC infections (especially by HIV or hepatitis B or C); inherited genetic
CC diseases; (auto)immune diseases (particularly rheumatism, arthritis,
CC arteriosclerosis, osteoporosis and diabetes, but many others listed) and
CC to prevent graft rejection. RH16 and its coding sequence are also useful
CC for inducing, or increasing, the immune response to a vaccine. The

CC present sequence was used in an example from the present invention
XX
SQ Sequence 1443 BP; 508 A; 261 C; 313 G; 360 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.: 6.57e-191 Length: 1443
Score: 2442.00 Matches: 475
Percent Similarity: 99.38% Conservative: 2
Best Local Similarity: 98.96% Mismatches: 3
Query Match: 45.98% Indels: 0
DB: 6 Gaps: 0

US-09-515-363C-2 (1-1025) x ABA04916 (1-1443)

QY	174	LysGluAsnTrpPheSerAlaPheLeuAsnValLeuArgGlnThrGlyAsnAsnGluLeu	193
Db	2	AAAGAAACTGGTTCCTGCAATTTCTGAATGTTCTTCGTCAAACAGGAAACAATGAAC	61
QY	194	ValGlnGluLeuThrGlySerAspCysSerGluSerAsnAlaGluIleGluAsnLeuSer	213
Db	62	GTCCAAGAGTTAACAGGCTCTGATTGCTCAGAAAGCAATGCAGAGATTGAGAA	121
QY	214	GlnValAspGlyProGlnValGluGlnLeuLeuSerThrThrValGlnProAsnLeu	233
Db	122	CAAGTTGATGGTCTCAAGTGGAGAGCAACTTCTTCAACCACAGTTCAGCCAAATCTG	181
QY	234	GluLysGluValTrpGlyMetGluAsnAsnSerSerGluSerSerPheAlaAspSerSer	253
Db	182	GAGAAGGAGGTCTGGGCATGGAGAATAACTCATCAGAAATCATCTTTTGCAGAT	241
QY	254	ValValSerGluSerAspThrSerLeuAlaGluGlySerValSerCysLeuAspGluSer	273
Db	242	GTAGTTTCAGAAATCAGACACAAGTTTGGCAGAAGGAAGTGTACGTCTTAGATGA	301
QY	274	LeuGlyHisAsnSerAsnMetGlySerAspSerGlyThrMetGlySerAspSerAspGlu	293
Db	302	CTTGGACATAACAGCAACATGGGCAGTGATTGAGGCACCATGGGAAGTGATTCA	361
QY	294	GluAsnValAlaAlaArgAlaSerProGluProGluLeuGlnLeuArgProTyrGlnMet	313
Db	362	GAGAAATGTGGCAGCAAGAGCATCCCGGAGCAGAACTCCAGCTCAGGCCTTACCA	421
QY	314	GluValAlaGlnProAlaLeuGluGlyLysAsnIleIleCysLeuProThrGlySer	333
Db	422	GAAAGTTGCCAGCCAGCCTTGGAGGGAAGAATATCATCATCTGCTCCCTACAGG	481
QY	334	GlyLysThrArgValAlaValTyrIleAlaLysAspHisLeuAspLysLysLysAla	353
Db	482	GGAAAACCCAGAGTGGCTGTTTACATTGCCAAGGATCACTTAGACAAGAAGAAAG	541
QY	354	SerGluProGlyLysValIleValLeuValAsnLysValLeuValGluGlnLeuPhe	373
Db	542	TCTGAGCCTGGAAAAGTTATAGTTCTTGTCATAAAGGTACTGCTAGTTGAACAG	601
QY	374	ArgLysGluPheGlnProPheLeuLysLysTyrTyrArgValIleGlyLeuSerGlyAsp	393
Db	602	CGCAAGGAGTTCCAACCATTTTGAAGAAATGGTATCGTGTATTGGATTAAAGTG	661
QY	394	ThrGlnLeuLysIleSerPheProGluValValLysSerCysAspIleIleSerThr	413
Db	662	ACCCAAGTGAATAATATCATTTCCAGAAGTGTCAAGTCCTGTGTATATTATCAG	721
QY	414	AlaGlnIleLeuGluAsnSerLeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGln	433
Db	722	GCTCAAATCCTTGAAAACCTCCCTCTTAAACTTGGAAAATGGAGAGATGCTGGT	781
QY	434	LeuSerAspPheSerLeuIleIleIleAspGluCysHisHisThrAsnLysGluAlaVal	453
Db	782	TTGTCAGACTTTTCCCTTCATTATCATTTGATGAATGTATCATCACCAACAAAG	841
QY	454	TyrAsnAsnIleMetArgHisTyrLeuMetGlnLysLeuLysAsnAsnArgLeuLys	473
Db	842	TATAATAACATCATGAGGCATTATTGATGCAGAAGTTTGAACAAATAGACTCAAG	901

QY	474	GluAsnLysProValIleProLeuProGlnIleLeuGlyLeuThrAlaSerProGlyVal	493
Db	902	GAACAACAAACAGTGATTTCCCTTCNTCAGATACTGGGACTAACAGCTTCACTGT	961
QY	494	GlyGlyAlaThrLysGlnAlaLysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeu	513
Db	962	GGAGGGGCCACGAAGCAAGCCAAAGCTGAAGAACACATTTTAAAACTATGTGCCA	1021
QY	514	AspAlaPheThrIleLysThrValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGln	533
Db	1022	GATGCATTTACTATTAAAACTGTTAAAGAAAAACCTTGATCAACTGAAAAACCA	1081
QY	534	GluProCysLysLysPheAlaIleAlaAspAlaThrArgGluAspPropPheLysGluLys	553
Db	1082	GAGCCATGCACAAGATTTGCCATTGCAGATGCAACCAGAGAAGATCCATTTAAAG	1141
QY	554	LeuLeuGluIleMetThrArgIleGlnThrTyrCysGlnMetSerProMetSerAspPhe	573
Db	1142	CTTCTAGAATAATGACAAGGATTCAAACTTATTGTCAATGAGTCCAATGTCAGAT	1201
QY	574	GlyThrGlnProTyrGluGlnTrpAlaIleGlnMetGluLysLysAlaAlaLysLysGly	593
Db	1202	GGAACTCAACCTATGAACAATGGCCATTCAAATGGAAAAAAGCTGCAAAAGAGGA	1261
QY	594	AsnArgLysGluArgValCysAlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIle	613
Db	1262	AATCGCAAGAAAGTGTTTGTGCAGAACATTTGAGGAAGTACAATAAGGCCCTAC	1321
QY	614	AsnAspThrIleArgMetIleAspAlaTyrThrHisLeuGluThrPheTyrAsnGlu	633
Db	1322	AATGACACAATTCGAATGATAGATGCGTATACTCATCTTGAACCTTCTATAATGA	1381
QY	634	LysAspLysLysPheAlaValIleGluAspAspSerAspGluGlyGlyAspAspGluTyr	653
Db	1382	AAAGATAAGAAGTTTGCAGTCATAGAAGATGATAGTGATGAGGGTGGTGATGAT	1441
RESULT 12			
ID	AAI59285	standard; cDNA; 1557 BP.	
XX	XX		
AC	AAI59285;		
XX	XX		
DT	22-OCT-2001	(first entry)	
XX	XX	Human polynucleotide SEQ ID NO 1488.	
DE	XX	Human; neurotropic; immunosuppressant; cytostatic; gene therapy; cancer;	
KW	XX	peripheral nervous system; neuropathy; central nervous system; CNS;	
KW	XX	Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;	
KW	XX	amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;	
KW	XX	chemokinetic; thrombolytic; drug screening; arthritis; inflammation;	
KW	XX	leukaemia; ss.	
OS	XX	Homo sapiens.	
XX	XX	WO200153312-A1.	
PN	XX	26-JUL-2001.	
PD	XX	26-DEC-2000; 2000WO-US034263.	
XX	XX	23-DEC-1999; 99US-00471275.	
PR	XX	21-JAN-2000; 2000US-00488725.	
PR	XX	25-APR-2000; 2000US-00552317.	
PR	XX	20-JUN-2000; 2000US-00598042.	
PR	XX	19-JUL-2000; 2000US-00620312.	
PR	XX	03-AUG-2000; 2000US-00653450.	
PR	XX	14-SEP-2000; 2000US-00662191.	
PR	XX	19-OCT-2000; 2000US-00693036.	
PR	XX	29-NOV-2000; 2000US-00727344.	
PA	XX	(HYSE-) HYSEQ INC.	

XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.

XX Homo sapiens.

XX WO200153312-A1.

XX 26-JUL-2001.

XX 26-DEC-2000; 2000WO-US034263.

XX 23-DEC-1999; 99US-00471275.

XX 21-JAN-2000; 2000US-00488725.

XX 25-APR-2000; 2000US-00552317.

XX 20-JUN-2000; 2000US-00598042.

XX 19-JUL-2000; 2000US-00620312.

XX 03-AUG-2000; 2000US-00653450.

XX 14-SEP-2000; 2000US-00662191.

XX 19-OCT-2000; 2000US-00693036.

XX 29-NOV-2000; 2000US-00727344.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;

XX Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;

XX Zhou P, Goodrich R, Drmanac RT;

XX WPI; 2001-442253/47.

XX P-PSDB; AAM41915.

XX Novel nucleic acids and polypeptides, useful for treating disorders such
PT as central nervous system injuries.
PT
XX
XX
XX Claim 1; SEQ ID NO 5060; 10078pp; English.
XX
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and the
CC encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification

XX
SQ Sequence 1382 BP; 513 A; 217 C; 285 G; 367 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.81e-164 Length: 1382
Score: 2120.00 Matches: 407
Percent Similarity: 98.32% Conservative: 2
Best Local Similarity: 97.84% Mismatches: 5
Query Match: 39.92% Indels: 2
DB: 4 Gaps: 0

US-09-515-363C-2 (1-1025) x AAI61071 (1-1382)

QY 612 GlnIleAsnAspThrIleArgMetIleAspAlaTyrThrHisLeuGluThrPheTyrAsn 631
Db |
13 CAAATTAATGACACAATTTCGAATGATAGATCGGTATCTCATCTTGAAACTTTCTATAAT 72

QY 632 GluGluLysAspLysLysPheAlaValIleGluAspAspSerAspGluGlyGlyAspAsp 651

Db |
73 GAAGAGAAAGATAAGAAGTTTGCAGTCATAGAAGATGATAGTGAGGGTGGTGATGAT 132
QY |
652 GluTyrCysAspGlyAspGluAspGluAspAspLeuLysLysProLeuLysLeuAspGlu 671
Db |
133 GAGTATTGTGATGGTGATGAACATGAGGATGATTTTAAGAAAACCTTTGAAAACCTGGATGAA 192
QY |
672 ThrAsp-ArgPheLeuMetThrLeuPhePheGluAsnAsnLysMet-LeuLysArgLeuA 691
Db |
193 ACAGATACGATTTCTCATGACTTTTATTTTGAACCAATAAATAGTTGAATAGGCTGG 252
QY |
691 laGluAsnProGluTyrGluAsnGluLysLeuThrLysLeuArgAsnThrIleMetGluG 711
Db |
253 CTGAACACCCCTTAATATGAAAATGAAAAGCTGACCAAAATTAAGAAAATACCATAATGGAGC 312
QY |
711 lnTyrThrArgThrGluGluSerAlaArgGlyIleIlePheThrLysThrArgGlnSerA 731
Db |
313 AATATACTAGGACTGAGGAATCAGCAGGAGGAATAATCTTTACAAAAACACGACAGAGTG 372
QY |
731 laTyrAlaLeuSerGlnTyrIleThrGluAsnGluLysPheAlaGluValGlyValIysA 751
Db |
373 CATATGCGCTTTCCAGTGGATTACTGAAAATGAAAATTTGCTGAAGTAGGAGTCAAAG 432
QY |
751 laHisHisLeuIleGlyAlaGlyHisSerSerGluPheLysProMetThrGlnAsnGluG 771
Db |
433 CCCACCATCTGATTGGAGCTGGACACAGCAGTGAGTTCAAACCCATGACACAGAAATGAAC 492
QY |
771 lnLysGluValIleSerLysPheArgThrGlyLysIleAsnLeuLeuIleAlaThrThrV 791
Db |
493 AAAAAGAAAGTCATTAGTAATTTCCGACTGGAAAAATAAATCTGCTTATCGCTACCACAG 552
QY |
791 alAlaGluGluGlyLeuAspIleLysGluCysAsnIleValIleArgTyrGlyLeuValT 811
Db |
553 TGGCAGAGAAGAGTCTGGATATTAAAGAATGTAACATTGTTATCCGTTATGGTCTCGTCA 612
QY |
811 hrAsnGluIleAlaMetValGlnAlaArgGlyArgAlaArgAlaAspGluSerThrTyrV 831
Db |
613 CCAATGAAATAGCCATGGTCCAGGCCGCTGGTCGAGCCAGAGCTGATGAGAGCACCTACG 672
QY |
831 alLeuValAlaHisSerGlySerGlyValIleGluHisGluThrValAsnAspPheArgG 851
Db |
673 TCCTGGTTGCTCACAGTGGTTTCAGGAGTTATCGAACATGACACAGATTAAATGATTTCCGAG 732
QY |
851 luLysMetMetTyrLysAlaIleHisCysValGlnAsnMetLysProGluGluTyrAlaH 871
Db |
733 AGAAGATGATGATAAAGCTATACATTGTGTTTCAAAATATGAAACCCAGAGGAGTATGCTC 792
QY |
871 isLysIleLeuGluLeuGlnMetGlnSerIleMetGluLysLysMetLysThrLysArgA 891
Db |
793 ATAAGATTTTGAATTTACAGATGCAAAAGTATAATGGAAGAAAATAATGAAAACCAAGAGAA 852
QY |
891 snIleAlaLysHisTyrLysAsnAsnProSerLeuIleThrPheLeuCysLysAsnCysS 911
Db |
853 ATATTGCCAAGCATTACAAGAATAACCCATCCTACTAATACTTTCTCTTGCAAAAACTGCA 912
QY |
911 erValLeuAlaCysSerGlyGluAspIleHisValIleGluLysMetHisHisValAsnM 931
Db |
913 GTGTGCTAGCCTGTTCTGGGGAAGATATCCATGTATTTGAGAAAATGCATCAGTCAATA 972
QY |
931 etThrProGluPheLysGluLeuTyrIleValArgGluAsnLysAlaLeuGlnLysLysC 951
Db |
973 TGACCCCAAGAAATTCAGGAACTTTACATTGTAAGAGAGAAAACAAAACACTGCAAAAGAAAGT 1032
QY |
951 ysAlaAspTyrGlnIleAsnGlyGluIleIleCysLysCysGlyGlnAlaTrpGlyThrM 971
Db |
1033 GTGCCGACTATCAAAATAAATGGTGAATAATCATCTGCAAAATGTGGCCAGGCTTGGGGAAACAA 1092
QY |
971 etMetValHisLysGlyLeuAspLeuProCysLeuLysIleArgAsnPheValValP 991
Db |
1093 TGATGGTGCACAAAGGCTTAGATTGCTCTGCTCAAAATAAGGAATTTTGTAGTGGTTT 1152
QY |
991 heLysAsnAsnSerThrLysLysGlnTyrLysLysTyrValGluLeuProIleThrPheP 1011

Db 1153 TCAAAATAAATTCAACAAAGAAACAATAACAAAAGTGGTAGAATTACTATCATCATTC 1212

QY 1011 roAsnLeuAspTyrSerGluCysCysLeuPheSerAspGluAsp 1025
|||||

Db 1213 CCAATCTTGACTATTTCAGAAATGCTGTTTATTATTAGTGATGAGGAT 1256

RESULT 14
ADC32572
ID ADC32572 standard; cDNA; 1382 BP.
XX
AC ADC32572;
XX
DT 18-DEC-2003 (first entry)
XX
DE Human novel cDNA contig sequence, SEQ ID NO:2654.
XX
KW Human; diagnostic; drug screening; forensics; gene mapping;
KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;
KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
KW ulcers; osteoporosis; autoimmune disease; cancer;
KW molecular weight marker; food supplement; antiparkinsonian; nootropic;
KW neuroprotective; antianaemic; anticoagulant; thrombolytic; vulnerary;
KW antiulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
KW gene therapy; ss.
XX
OS Homo sapiens.
XX
PN WO2003029271-A2.
XX
PD 10-APR-2003.
XX
XX 24-SEP-2002; 2002WO-US030474.
PF
XX 24-SEP-2001; 2001US-0324631P.
PR
XX (HYSE-) HYSEQ INC.
XX
XX Tang TV, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;
PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
PI Haley-Vicente D, Drmanac RT;
XX
XX WPI; 2003-371981/35.
DR P-PSDB; ADC33339.
XX
XX New polynucleotide and polypeptide useful for diagnosing, preventing or
PT treating conditions such as neurodegenerative diseases, anemias, platelet
PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
PT cancer.
XX
PS Example 2; SEQ ID NO 2654; 1185pp; English.
XX
XX The invention relates to 971 novel human cDNA sequences (ADC29919-
CC ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The
CC invention also relates to nucleic acid sequences over 99% identical with
CC the novel human cDNAs. The invention additionally encompasses expression
CC vectors and host cells comprising a nucleic acid of the invention; the
CC recombinant production of a polypeptide of the invention; an antibody
CC against a polypeptide of the invention; a method of detecting
CC polynucleotides or polypeptides of the invention; and methods of
CC identifying a compound which binds to a polypeptide of the invention. The
CC invention further discloses methods of preventing, treating or
CC ameliorating a medical condition; kits comprising polynucleotide probes
CC and/or monoclonal antibodies for carrying out the methods of the
CC invention; methods for the identification of compounds that modulate the
CC expression or activity of the polynucleotide and/or polypeptide; and 767
CC contig sequences corresponding to the cDNA sequences of the invention
CC (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628
CC -ADC33394). The nucleic acids and polypeptides of the invention are
CC useful in diagnostics, drug screening, forensics, gene mapping, in the
CC identification of mutations responsible for genetic disorders or other
CC traits, for assessing biodiversity, and in producing many other types of
CC data and products dependent on DNA and amino acid sequences. They are
CC also used for treating diseases such as Parkinson's disease, Alzheimer's

CC disease and other neurodegenerative diseases, anaemia, platelet
CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
CC cancer. The nucleic acids may also be used as hybridisation probes or
CC primers, and in the recombinant production of a protein. The polypeptides
CC are also useful in generating antibodies, as molecular weight markers,
CC and as food supplements. The present sequence represents a human contig
CC sequence used in an example of the invention. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1382 BP; 513 A; 217 C; 285 G; 367 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.81e-164 Length: 1382
Score: 2120.00 Matches: 407
Percent Similarity: 98.32% Conservative: 2
Best Local Similarity: 97.84% Mismatches: 5
Query Match: 39.92% Indels: 2
DB: 10 Gaps: 0

US-09-515-363C-2 (1-1025) x ADC32572 (1-1382)
QY 612 GlnIleAsnAspThrIleArgMetIleAspAlaTyrThrHisLeuGluThrPheTyrAsn 631
Db 13 CAAATTAATGACACAATTCGAATGATAGATGCGTATATCTCATCTTGAAACTTCTATAAT 72
QY 632 GluGluLysAspLysLysPheAlaValIleGluAspAspSerAspGluGlyGlyAspAsp 651
Db 73 GAAGAGAAAGATAAAGATTTTGCAGTCATAGAAAGATGATAGTGGGTGGTGATGAT 132
QY 652 GluTyrCysAspGlyAspGluAspGluAspLeuLysLysProLeuLysLeuAspGlu 671
Db 133 GAGTATTGTGATGCGTATGAACATGAGGATGATTTTAAGAAACCTTTGAAACTGGATGAA 192
QY 672 ThrAsp-ArgPheLeuMetThrLeuPhePheGluAsnAsnLysMet-LeuLysArgLeuA 691
Db 193 ACAGATACGATTTCTCATGACTTTATTTTGGAAAACAATAAAATAGTTGAATAGGCTGG 252
QY 691 laGluAsnProGluTyrGluAsnGluLysLeuThrLysLeuArgAsnThrIleMetGluG 711
Db 253 CTGAACACCCCTAATATGAAATGAAAGCTGACCAATTAAGAAATACCATATATGGAGC 312
QY 711 lntyrThrArgThrGluGluSerAlaArgGlyIleIlePheThrLysThrArgGlnSerA 731
Db 313 AATATACTAGGACTGAGGAATCAGCACGAGGAATAATCTTTACAAAAAACACGACAGAGTG 372
QY 731 laTyrAlaLeuSerGlnTrpIleThrGluAsnGluLysPheAlaGluValGlyValLysA 751
Db 373 CATATGCGCTTTCCAGTGGATTACTGAAAATGAAAAATTTTGCTGAAGTAGGAGTCAAAG 432
QY 751 laHisHisLeuIleGlyAlaGlyHisSerSerGluPheLysProMetThrGlnAsnGluG 771
Db 433 CCCACCATCTGATTGGAGCTGGACACAGCAGTGAGTTCAAACCCATGACACAGAATGAAC 492
QY 771 lnlLysGluValIleSerLysPheArgThrGlyLysIleAsnLeuLeuIleAlaThrThrV 791
Db 493 AAAAAGAAGTCATTAGTAAATTTGCACTGGAAAAATAAATCTGCTTATCGTACCACAG 552
QY 791 alAlaGluGluGlyLeuAspIleLysGluCysAsnIleValIleArgTyrGlyLeuValT 811
Db 553 TGGCAGAAGAAGGTCCTGGATATTAAAGAAATGTAACATTTGTTATCCGTTATGGTCTCGTCA 612
QY 811 hrAsnGluIleAlaMetValGlnAlaArgGlyArgAlaArgAlaAspGluSerThrTyrV 831
Db 613 CCAATGAATAGCCATGGTCCAGGCCCGTGGTCGAGCCAGAGCTGATGAGAGCACCTACG 672
QY 831 alLeuValAlaHisSerGlySerGlyValIleGluHisGluThrValAsnAspPheArgG 851
Db 673 TCCTGGTTCACAGTGGTTTCAGGAGTTATCGAACATGAGACAGTGAATGATTCGAG 732
QY 851 luLysMetMetTyrLysAlaIleHisCysValGlnAsnMetLysProGluGluTyrAlaH 871

Db 733 AGAAGATGATGTATAAAGCTATACATTGTGTCTCAAATATGAACCAGAGGATGCTC 792

QY isLysIleLeuGluGlnMetGlnSerIleMetGluLysMetLysThrLysArgA 891

Db 793 ATAAGATTTTGGAAATTACAGATGCAAAAGTATAATGGAAGAAATGAAACCAAGAGAA 852

QY snIleAlaLysHisTyrLysAsnAsnProSerLeuIleThrPheLeuCysLysAsnCys 911

Db 853 ATATTGCCAAGCATTACAAGAATAACCCATCACTAATAACTTTCCTTGCAAAACTGCA 912

QY erValLeuAlaCysSerGlyGluAspIleHisValIleGluLysMetHisHisValAsnM 931

Db 913 GTGTGCTAGCCTGTTCTGGGAAGATATCCATGTAAATTGAGAAATGCATCACGTCAATA 972

QY etThrProGluPheLysGluLeuTyrIleValArgGluAsnLysAlaLeuGlnLysLysC 951

Db 973 TGACCCAGAAATTCAGGAACCTTTACATTGTAAGAGAAACAAACACTGCAAAAGAGT 1032

QY ysAlaAspTyrGlnIleAsnGlyGluIleIleCysLysCysGlyGlnAlaTrpGlyThrM 971

Db 1033 GTGCCGACTATCAATAAATGGTGAAATCATCTGCAAAATGTGCCAGGCTTGGGGAACAA 1092

QY etMetValHisLysGlyLeuAspLeuProCysLeuLysIleArgAsnPheValValP 991

Db 1093 TGATGGTGCACAAAGCCTTAGATTTGCCTTGCTCAAAATAAGCAATTTTGTAGTGGTTT 1152

QY heLysAsnAsnSerThrLysLysGlnTyrLysLysTyrValGluLeuProIleThrPheP 1011

Db 1153 TCAAAAATAATTCAACAAAGAAACAATAACAAAAAGTGGGTAGATTACCTATCACATTTC 1212

QY roAsnLeuAspTyrSerGluCysCysLeuPheSerAspGluAsp 1025

Db 1213 CCAATCTTGACTATTCAGAATGCTGTTTATTATTAGTGATGAGGAT 1256

RESULT 15

ADI30722

ID ADI30722 standard; cDNA; 1392 BP.

XX

AC ADI30722;

XX

DT 17-JUN-2004 (first entry)

XX

DE Human cDNA #48.

XX

KW Human; gene; ss; immunological response; immunopathological condition; Crohn's disease; asthma; ulcerative colitis; hypereosinophilia; irritable bowel syndrome; osteoarthritis; rheumatoid arthritis; acute monocytic leukaemia; antiinflammatory; antiasthmatic; antiulcer; osteopathic; antiarthritic; antirheumatic; cytostatic.

XX

OS Homo sapiens.

XX

PN US6607879-B1.

XX

PD 19-AUG-2003.

XX

PF 09-FEB-1998; 98US-00023655.

XX

PR 09-FEB-1998; 98US-00023655.

XX

PA (INCY-) INCYTE CORP.

XX

PI Cocks BG, Stuart SG, Seilhamer JJ;

XX

DR WPI; 2003-895307/82.

XX

PT A composition comprising a plurality of cDNAs, useful for detecting altered expression of genes in an immunological response or for diagnosing and treating an immunopathology, e.g. Crohn's disease, asthma or osteoarthritis.

PT

XX

PS Claim 1; SEQ ID NO 48; 50pp; English.

XX

CC The invention relates to a composition comprising a plurality of cDNAs for detecting the altered expression of genes in an immunological response. The invention also relates to a method of diagnosing or monitoring the treatment of an immunopathological condition in a sample, comprising obtaining nucleic acids from a sample, contacting the nucleic acids of the sample with an array comprising the plurality of cDNAs under conditions to form one or more hybridisation complexes, detecting the hybridisation complexes and comparing the levels of hybridisation complexes detected in a non-diseased sample, where an altered level of the detected hybridisation complexes correlates with the presence of an immunopathological condition. Also disclosed are an expression profile comprising a microarray and a plurality of detectable complexes and a method for identifying a plurality of polynucleotide probes. The cDNAs are useful as hybridisable array elements in a microarray for monitoring the expression of target polynucleotides. The microarray can be used in the diagnosis of an immunopathology, such as Crohn's disease, asthma, ulcerative colitis, hypereosinophilia, irritable bowel syndrome, osteoarthritis, rheumatoid arthritis or acute monocytic leukaemia, and in identifying agents for the treatment of the diseases. The microarray may also be used in drug discovery and development, toxicological and carcinogenicity studies, forensics or pharmacogenomics. The composition may also be used in purification of a subpopulation of mRNAs, cDNAs or genomic fragments. This sequence represents a human cDNA of the invention. Note: The sequence data for this patent did not form part of the printed specification but was obtained in electronic format directly from USPTO at seqdata.uspto.gov/sequence.html.

XX

SQ Sequence 1392 BP; 497 A; 215 C; 264 G; 351 T; 0 U; 65 Other;

Alignment Scores:

Pred. No.: 1.61e-160 Length: 1392

Score: 2072.00 Matches: 400

Percent Similarity: 94.79% Conservative: 0

Best Local Similarity: 94.79% Mismatches: 22

Query Match: 39.01% Indels: 0

DB: 11 Gaps: 0

US-09-515-363C-2 (1-1025) x ADI30722 (1-1392)

QY 604 LeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMetIleAspAlaTyr 623

Db 3 TTGAGGAAGTACAATGAGGCCCTACAAATTAAATGACACAAATTCGAATGATAGATCGGTAT 62

QY 624 ThrHisLeuGluThrPheTyrAsnGluGluLysAspLysPheAlaValIleGluAsp 643

Db 63 ACTCATCTTGAAACTTTCTATAATGAAGAGAAAGATAAGAAGTTTGCAATANNNNNN 122

QY 644 AspSerAspGluGlyGlyAspAspGluTyrCysAspGlyAspGluAspGluAspLeu 663

Db 123 NNTTA 182

QY 664 LysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeuPhePheGluAsn 683

Db 183 AAGAAACCTTTGAAACTGGATGAAACAGATAGATTTCTCATGACTTTATTTTGGAAAC 242

QY 684 AsnLysMetLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGluLysLeuThrLys 703

Db 243 AATAAAATGTTGAAAAGGCTGGCTGAAAACCCAGAAATATGAAAATGAAAAGCTGACCAA 302

QY 704 LeuArgAsnThrIleMetGluGlnTyrThrArgThrGluGluSerAlaArgGlyIleIle 723

Db 303 TTAAGAAATACCATATATGGAGCAATATACTAGGACTGAGGAATCAGCACGAGGAATATC 362

QY 724 PheThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTrpIleThrGluAsnGluLys 743

Db 363 TTTACAAAACACGACAGAGTGCATATGCGCTTTCCAGTGGATTACTGAAAATGAAAAA 422

QY 744 PheAlaGluValGlyValLysAlaHisHisLeuIleGlyAlaGlyHisSerSerGluPhe 763

Db 423 TTTGCTGAAGTAGGAGTCAAAGCCCACTCTGATTGGAGCTGGACACAGCAGTGAGTTC 482

QY 764 LysProMetThrGlnAsnGluGlnLysGluValIleSerLysPheArgThrGlyLysIle 783

Db		483	AAACCCATGACACAGAAATGAACAAAAGAGTCATTAGTAAATTTTCGCACCTGGAAAAATA	542
QY		784	AsnLeuLeuIleAlaThrThrValAlaGluGluGlyLeuAspIleLysGluCysAsnIle	803
Db		543	AATCTGCTTATCGCTACCCACAGTGGCAGAGAAGGCTGGATATTAAGAATGTAAACATT	602
QY		804	ValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAlaArgGlyArgAla	823
Db		603	GTATCCGTTATGGTCTCGTCACCAATGAATAGCCATGGTCCAGGCCCGGTGGTCGAGCC	662
QY		824	ArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGlyValIleGluHis	843
Db		663	AGAGCTGATGAGAGCACCTACGTCCTGGTTGCTCACAGTGGTTCAGGAGTTATCGAACGT	722
QY		844	GluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHisCysValGlnAsn	863
Db		723	GAGACAGTTAATGATTTCCGAGAGAGAGATGATGTATAAGCTATACATTGTGTCAAAAT	782
QY		864	MetLysProGluGluTyrAlaHisLysIleLeuGluLeuGlnMetGlnSerIleMetGlu	883
Db		783	ATGAACCCAGAGAGATGCTCATAGATTTTGGAAATTACAGATGCAAAAGTATAATGGAA	842
QY		884	LysLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAsnProSerLeuIle	903
Db		843	AAGAAAATGAAAACCAAGAGAAATATTGCCAAGCATTACAAGATAACCCATCACTAATA	902
QY		904	ThrPheLeuCysLysAsnCysSerValLeuAlaCysSerGlyGluAspIleHisValIle	923
Db		903	ACTTTCCCTTTGCAAAAAAAGCTGCAGTGTGTAGCCTGTTCTGGGGAAGATATCCATGTAAT	962
QY		924	GluLysMetHisHisValAsnMetThrProGluPheLysGluLeuTyrIleValArgGlu	943
Db		963	GAGAAAATGCATCACGTCATATATGACCCAGAAATTCAGGAACTTTACATTGTAAAGAGAA	1022
QY		944	AsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluIleIleCysLys	963
Db		1023	AACAAAGCACTGCAAAAGAAAGTGTGCCGACTATCAAAATAAATGGTGAATCATCTGCAAA	1082
QY		964	CysGlyGlnAlaTrpGlyThrMetMetValHisLysGlyLeuAspLeuProCysLeuLys	983
Db		1083	TGTGGCCAGGCTTGGGGAACAATGATGGTGCAAAAGGCTTAGATTTGCCTTGTCTCAAA	1142
QY		984	IleArgAsnPheValValPheLysAsnAsnSerThrLysLysGlnTyrLysLysTrp	1003
Db		1143	ATAAGGAATTTGTAGTGGTTTCAAAAATAATTCAACAAAGAAACAATACAAAAAGTGG	1202
QY		1004	ValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCysCysLeuPheSerAsp	1023
Db		1203	GTAGAATTACCTATACATTTCCCAATCTTGACTATTTCAGAATGCTGTTATTATTAGTGAT	1262
QY		1024	GluAsp	1025
Db		1263	GAGGAT	1268

TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 1392 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: HMCINOT01
CLONE: 004700
US-09-023-655-48

Alignment Scores:

Pred. No.: 9.26e-229 Length: 1392
Score: 2072.00 Matches: 400
Percent Similarity: 94.79% Conservative: 0
Best Local Similarity: 94.79% Mismatches: 22
Query Match: 39.01% Indels: 0
DB: 4 Gaps: 0

US-09-515-363C-2 (1-1025) x US-09-023-655-48 (1-1392)

QY 604 LeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMetIleAspAlaTyr 623
Db 3 TTGAGGAAGTACAAATGAGGCCCTACAAATTAATGACACAAATTCGAATGATAGATGCGGTAT 62
QY 624 ThrHisLeuGluThrPheTyrAsnGluLysAspLysPheAlaValIleGluAsp 643
Db 63 ACTCATCTTGAAACTTCTATAATGAAGAGAAAGATAAGAAGTTTGACGTCATANNNNN 122
QY 644 AspSerAspGluGlyGlyAspAspGluTyrCysAspGlyAspGluAspGluAspLeu 663
Db 123 NNTTA 182
QY 664 LysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeuPhePheGluAsn 683
Db 183 AAGAAACCTTGAAACTGGATGAACAGATAGATTTCTCATGACTTTATTTTTTGAAAC 242
QY 684 AsnLysMetLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGluLysLeuThrLys 703
Db 243 AATAAAATGTTGAAAGGGCTGGCTGAAACCCAGAAATATGAAATGAAAGCTGACCAA 302
QY 704 LeuArgAsnThrIleMetGluGlnTyrThrArgThrGluSerAlaArgGlyIleIle 723
Db 303 TTAAGAAATACCATAAATGGAGCAATATACCTAGGACTGAGGAATCAGCACGAGGAATC 362
QY 724 PheThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTrpIleThrGluAsnGluLys 743
Db 363 TTTACAAAAACACGACAGAGTGATATGCGCTTTCCAGTGGATTACTGAAATGAAAAA 422
QY 744 PheAlaGluValGlyValLysAlaHisHisLeuIleGlyAlaGlyHisSerSerGluPhe 763
Db 423 TTTGCTGAAGTAGGAGTCAAAGCCCCACCATCTGATTGGAGTGGACACAGCAGTGAGTTC 482
QY 764 LysProMetThrGlnAsnGluGlnLysGluValIleSerLysPheArgThrGlyLysIle 783
Db 483 AAACCCATGACACAGAAATGAACAAAAGAAAGTCATTAGTAAATTTCCGACTGGAAAAA 542
QY 784 AsnLeuLeuIleAlaThrThrValAlaGluGluGlyLeuAspIleLysGluCysAsnIle 803
Db 543 AATCTGCTTATCGCTACCACAGTGGCAGAAAGGCTCTGGATATTAAAGAAATGTAACATT 602
QY 804 ValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAlaArgGlyArgAla 823
Db 603 GTTATCCGTTATGGTCTCGTCCACCAATGAAATAGCCATGGTCCAGGCCCGTGGTCGAGCC 662
QY 824 ArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGlyValIleGluHis 843
Db 663 AGAGCTGATGAGACACCTACGTCCTGGTGTCTCACAGTGGTTCAGGAGTTATCGAACGT 722
QY 844 GluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHisCysValGlnAsn 863

Db 723 GAGACAGTTAATGATTTCCGAGAGAAAGATGATGTATATAAGCTATACATTTGTTCAAAAT 782
QY 864 MetLysProGluGluTyrAlaHisLysIleLeuGluLeuGlnMetGlnSerIleMetGlu 883
Db 783 ATGAAACACGAGGAGTATGCTCATAGATTTTGAATTTACAGATGCAAAAGTATAATGGAA 842
QY 884 LysLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAsnProSerLeuIle 903
Db 843 AAGAAAATGAAACCAAGAGAAATATTGCCAAGCATTACAAGAATAAACCCTACTAATA 902
QY 904 ThrPheLeuCysLysAsnCysSerValLeuAlaCysSerGlyGluAspIleHisValIle 923
Db 903 ACTTTCCTTTGCAAAAACCTGCAGTGTGCTAGCTGTTCTGGGGAAGATATCCATGTAATT 962
QY 924 GluLysMetHisHisValAsnMetThrProGluPheLysGluLeuTyrIleValArgGlu 943
Db 963 GAGAAAATGCAATCACGTCAATATGACCCCAAGAAATTCAGGAACCTTTACATTTGTAAGAGAA 1022
QY 944 AsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluIleCysLys 963
Db 1023 AACAAAGCACTGCAAAAAGAAAGTGTGCCGACTATCAAAATAAATGGTGAATCATCTGCAAA 1082
QY 964 CysGlyGlnAlaTyrGlyThrMetMetValHisLysGlyLeuAspLeuProCysLys 983
Db 1083 TGTGGCCAGGCTTGGGGAACAATGATGTTGTCACAAAGGCTTAGATTTGCTCTCAAA 1142
QY 984 IleArgAsnPheValValPheLysAsnAsnSerThrLysLysGlnTyrLysLysTrp 1003
Db 1143 ATAAGGAATTTTGTAGTGGTTTTCAAAAATAATTCACAAAGAAACATACAAAAGTGG 1202
QY 1004 ValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCysCysLeuPheSerAsp 1023
Db 1203 GTAGAATTACCTATACATTTCCCAATCTTGACTATTTCAGAAATGCTGTTATTAGTGAT 1262
QY 1024 GluAsp 1025
Db 1263 GAGGAT 1268
RESULT 2
US-09-799-451-771
; Sequence 771, Application US/09799451
; Patent No. 6783969
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie
; APPLICANT: Xue, Aidong J.
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Ma, Yungqing
; APPLICANT: Yamazaki, Victoria
; APPLICANT: Chen, Rui-hong
; APPLICANT: Wang, Zhiwei
; APPLICANT: Wang, Dunrui
; APPLICANT: Yang, Yonghong
; APPLICANT: Wehrman, Tom
; APPLICANT: Ghosh, Reena
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6783969el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 803
; CURRENT APPLICATION NUMBER: US/09/799,451
; CURRENT FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 948
; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 771
; LENGTH: 2590
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:

QY 964 ---CysGlyGlnAlaTrpGlyThrMetMetValHisLysGlyLeuAspLeuProCysLeu 982
Db 2029 AACTGTGGGAGGTCTGGGGTCTGCAGATGATCTACAAGTCAGTGAAGCTGCCAGTGCTC 2088
QY 983 LysIleArgAsnPheValValPheLysAsnAsnSerThrLysLysGlnTyrLysLys 1002
Db 2089 AAAGTCCGCAGC-----ATGCTGCTGGAGACCCCTCAGGGCGGATCCAGGCCAAAAG 2142
QY 1003 TrpValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCysCys 1019
Db 2143 TGGTCCCGGTGGCCCTTCTCCGTGCCTGACTTTGACTTCTCTGCAGCATTGT 2193
RESULT 3
US-08-143-576-7
; Sequence 7, Application US/08143576
; Patent No. 5643761
; GENERAL INFORMATION:
; APPLICANT: Fisher, Paul B.
; APPLICANT: Jiang, Hongping
; TITLE OF INVENTION: METHOD FOR GENERATING A SUBTRACTED
; TITLE OF INVENTION: CDNA LIBRARY AND USES OF THE GENERATED LIBRARY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John P. White, c/o Cooper & Dunham
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/143,576
; FILING DATE: 25-OCT-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 43563/JPW/AKC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 977-9550
; TELEFAX: (212) 664-0525
; TELEX: 422523 COOP UI
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 301 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-143-576-7

Alignment Scores:
Pred. No.: 3.49e-45 Length: 301
Score: 472.00 Matches: 97
Percent Similarity: 97.98% Conservative: 0
Best Local Similarity: 97.98% Mismatches: 2
Query Match: 8.89% Indels: 1
DB: 1

US-09-515-363C-2 (1-1025) x US-08-143-576-7 (1-301)

QY 476 LysProValIleProLeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyGly 495
Db 3 AAACCAAGTGATCCCTTCCTCAGATACTGGGACTAACAGCTTCACCTGGTGTGGAGGG 62
QY 496 AlaThrLysGlnAlaLysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeuAspAla 515

Db 63 GCCACGAAGCAAGCCAAAGCTGAAGAACACATTTTAAAACTATGTGCCTATCTTGATGCA 122
QY 516 PheThrIleLysThrValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluPro 535
Db 123 TTTTACTATTAAAACTGTTAAAGAAAAACCTTGATCAACTGAAAAACCAAAATACAGGAGCA- 181
QY 536 CysLysLysPheAlaIleAlaAspAlaThrArgGluAspPropheLysGluLysLeuLeu 555
Db 182 TGCAAGAAGTTTGCCATTGCAGATGCAACCAGAGAAGATCCATTTTAAAGAGAAAACTTCTA 241
QY 556 GluIleMetThrArgIleGlnThrTyrCysGlnMetSerProMetSerAspPheGly 574
Db 242 GAAATAATGACAAGGATTCAAACCTTATTGTCAAATGAGTCCAATGTCAGATTTTTGA 298

RESULT 4

US-09-221-268D-12
; Sequence 12, Application US/09221268D
; Patent No. 6720408
; GENERAL INFORMATION:
; APPLICANT: Fisher, Paul B.
; APPLICANT: Jiang, Hongping
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES WHICH SPECIFICALLY HYBRIDIZE WITH
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCE ENCODING MDA-7 AND PHARMACEUTICAL COMPOSI
; FILE REFERENCE: A34534-A-A (070050.1637)
; CURRENT APPLICATION NUMBER: US/09/221,268D
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: 08/316,537
; PRIOR FILING DATE: 1994-09-30
; PRIOR APPLICATION NUMBER: 08/143,576
; PRIOR FILING DATE: 1993-10-27
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 301
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-221-268D-12

Alignment Scores:
Pred. No.: 3.49e-45 Length: 301
Score: 472.00 Matches: 97
Percent Similarity: 97.98% Conservative: 0
Best Local Similarity: 97.98% Mismatches: 2
Query Match: 8.89% Indels: 1
DB: 4

US-09-515-363C-2 (1-1025) x US-09-221-268D-12 (1-301)

QY 476 LysProValIleProLeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyGly 495
Db 3 AAACCAAGTGATCCCTTCCTCAGATACTGGGACTAACAGCTTCACCTGGTGTGGAGGG 62
QY 496 AlaThrLysGlnAlaLysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeuAspAla 515
Db 63 GCCACGAAGCAAGCCAAAGCTGAAGAACACATTTTAAAACTATGTGCCTATCTTGATGCA 122
QY 516 PheThrIleLysThrValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluPro 535
Db 123 TTTTACTATTAAAACTGTTAAAGAAAAACCTTGATCAACTGAAAAACCAAAATACAGGAGCA- 181
QY 536 CysLysLysPheAlaIleAlaAspAlaThrArgGluAspPropheLysGluLysLeuLeu 555
Db 182 TGCAAGAAGTTTGCCATTGCAGATGCAACCAGAGAAGATCCATTTTAAAGAGAAAACTTCTA 241
QY 556 GluIleMetThrArgIleGlnThrTyrCysGlnMetSerProMetSerAspPheGly 574
Db 242 GAAATAATGACAAGGATTCAAACCTTATTGTCAAATGAGTCCAATGTCAGATTTTTGA 298

RESULT 5

US-09-590-968B-1
; Sequence 1, Application US/09590968B
; Patent No. 6737561
; GENERAL INFORMATION:

```
; APPLICANT: Ray, Animesh
; APPLICANT: Golden, Teresa Ann
; TITLE OF INVENTION: GENE ENCODING SHORT INTEGUMENTS AND USES THEREOF
; FILE REFERENCE: 176/60581
; CURRENT APPLICATION NUMBER: US/09/590,968B
; CURRENT FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/138,316
; PRIOR FILING DATE: 1999-06-09
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 6184
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; US-09-590-968B-1

Alignment Scores:
Pred. No.:      3.25e-35      Length:      6184
Score:          406.00      Matches:      231
Percent Similarity: 39.82%      Conservative: 158
Best Local Similarity: 23.64%      Mismatches:  316
Query Match:     7.64%      Indels:      275
DB:              4          Gaps:         47

US-09-515-363C-2 (1-1025) x US-09-590-968B-1 (1-6184)

QY 20 AlaArgValLysMetTyrIleGlnValGluProValLeuAspTyrLeuThrPheLeuPro 39
Db 127 TCTAGGGTTTGGCTGCTCTCTTCTCGTGACCC-----TTTTTACCT 168

QY 40 Ala-----GluValLysGluGlnIleGlnArgThrValAlaThrSerGlyAsnMetGln 57
Db 169 GCAACAACAACATTCAAAATTGGCGTGTTCGTACGGTCTA-TCTAACCCCTAATCTGTCA 227

QY 58 AlaValGluLeuLeuSerThrLeuGluLysGlyValTrpHisLeuGlyTrpThrArg 77
Db 228 CAAACACACTCTCTCTCACCCCTTTTCTGGGTTTATTCAAT-----TCTCGT 278

QY 78 GluPheValGluAlaLeuArgArgThrGlySerProLeuAlaAlaArgTyrMetAsnPro 97
Db 279 GCTTTT-----GGTCTGTTTTTCTTCTCTGGG 305

QY 98 GluLeuThrAspLeuProSerProSerPheGluAsnAla-----HisAspGlu 113
Db 306 GATTGGTTTTCTTG---AGTGAGTTTTTCTCCTCTTTCTTTATGTTCTTGATTGATTAT 362

QY 114 TyrLeuGlnLeu-----LeuAsnLeuLeuGlnPro-----ThrLeuVal 126
Db 363 TATATAGAATTATGTAATGGAGGATGAGCTTAGAGAGCCACCAATAAAGCCTTCITTATT 422

QY 127 AspLysLeuLeuValArgAspValLeuAspLysCys---MetGluGluGluLeuThr 145
Db 423 GGCTAGATGCTTGC-GAGGACATC-----TCTTGATCTTATCGATGATCTCGTGCT 475

QY 146 IleGluAspArgAsnArgIleAlaAlaAlaGluAsnAsnGlyAsnGluSerGlyValArg 165
Db 476 GAATTTGATCCTTCTCTGTTGCTGTCATGAATCCACT---GATGAAAACGGCGTC--- 529

QY 166 GluLeuLysArgIleValGlnLysGluAsnTrpPheSerAlaPheLeuAsnValLeu 185
Db 530 -----ATCAATGATTTTTTCGGTGGGATTGATCACATTTTA 565

QY 186 ArgGlnThrGlyAsnAsnGluLeuValGlnGluLeuThrGlySerAspCysSerGluSer 205
Db 566 GATAGTATCAAGAACGGT-----GGAGGCTTACCAACCAATGGCGTTTCTGTATACC 616

QY 206 AsnAlaGluIleGluAsnLeuSerGlnValAspGlyProGlnValGluGlnLeuLeu 225
Db 617 AATTCTCAAATCAACGAGGTACTGTAACT-----CCTCAGGT-----ATTGCT 661

QY 226 SerThrThrValGlnProAsn---LeuGluLysGluValTrpGlyMetGluAsnAsnSer 244
Db 662 AAGGAGACAGTGAAGGAGAATGGGTTGCAAAAAGAATGGCGGTAAGAGAGACGAATTCTCG 721
```

```
QY 245 SerGluSerSerPheAlaAsp-----SerSerValValSerGluSer 258
Db 722 AAAGAGGAGGAGACAAGGATAGGAAGAGAGCTAGGGTTTGTAGTTATCAGAGTGAAGG 781

QY 259 AspThrSerLeuAlaGluGlySerValSerCysLeuAspGlu----- 272
Db 782 AGTAACCTTTCAGGTAGAGGGCATGTTAATAATTCTAGGAGGGAGATAGTTTATGAAT 841

QY 273 -----SerLeuGlyHisAsn----- 277
Db 842 AGGAAACGTACTCGTAATTGGGACGAGCGGGTAACAATAAGAAAGGAATGTAAC 901

QY 278 -----SerAsnMet 280
Db 902 AATTACAGAAGAGATGGTAGAGATAGAGAAAGTTAGGGTTATTGGGAGGGGATAAAGTT 961

QY 281 GlySerAsp-----SerGlyThrMetGlySerAspSerAsp----- 292
Db 962 GGTTCCAATGAGTTGGTTTATAGGTCAGGGACTTGGGAAGCTGATCATGAAAGAGATGTT 1021

QY 293 -----GluGluAsnValAlaAla 298
Db 1022 AAGAAAGTCAGTGGTGGAAACCCGGAATGCGATGTCAAGGCAGAGGAGAAC----- 1072

QY 299 ArgAlaSerProGluPro-----GluLeuGlnLeuArgProTyrGln 312
Db 1073 AAGAGTAAGCCTGAAGAACGTAAGAGAAAGGTTGTGGAGAGCAAGCAAGCGATACCAG 1132

QY 313 MetGluValAlaGlnProAlaLeuGluGlyLysAsnIleIleCysLeuProThrGly 332
Db 1133 TTGGATGTTCTTGAACAAGCT---AAAGCGAAAACACGATGCTTCTCTTGAGACCGGT 1189

QY 333 SerGlyLysThrArgValAlaValTyr-----IleAlaLysAspHisLeuAsp 348
Db 1190 GCTGGAAGACACTTATCGCGATTCTTCTTATTAAAGTGTTCATAAGGATCTGATGAGC 1249

QY 349 LysLysLysLysAlaSerGluProGlyLysValIleValLeuValAsnLysValLeuLeu 368
Db 1250 CAGAACAGAAAATGCTCTCG-----GTGTTCTTGGTTCCCAAAGTGCCCTTG 1297

QY 369 ValGluGlnLeuPheArgLysGlu-----PheGlnProPheLeuLysLys 383
Db 1298 GTTTATCAGCA-AGCAGAAAGTGATCCGTAATCAAACCTTGTTTCAAGTTGGACATTATTG 1356

QY 384 TrpTyrArgValIleGlyLeuSerGlyAspThrGlnLeu-LysIleSerPheProGluVa 403
Db 1357 TGGTGAGATGGGACAGGACTTTTGGGATTTCTCGAAGGTGGCAACGAGAGTTT----- 1408

QY 403 lValLysSerCysAspIleIleIleSerThrAlaGlnIleLeuGluAsnSerLeuLeuAs 423
Db 1409 ----GAGTCTAAGCAGGTTCTAGTTATGACAGCACAAATCTGTTGAATATACTG---- 1459

QY 423 nLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPheSerLeuIleIleAs 443
Db 1460 -----AGACACAGTATCATTAGAATGGAACAACAAATTGATCTTCTTATTCTCGA 1506

QY 443 pGluCysHisHisThrAsnLysGluAlaValTyrAsnAsnIleMetArgHisTyrLeuMe 463
Db 1507 CGAGTGTCACCACGCTGTCAAGAAACATCCATACTCTTTAGTGTGTCAGAGTTTTACCA 1566

QY 463 tGlnLysLeuLysAsnAsnArgLeuLysLysGluAsnLysProValIleProLeuProGl 483
Db 1567 TACAACTCCTAAAAGATAAAAGA-----CCTGC 1593

QY 483 nIleLeuGlyLeuThrAlaSerPro---GlyValGlyGlyAlaThrLysGlnAlaLysAl 502
Db 1594 CATCTTTGGAATGACTGCTTCGCTGTAAATTAAAGSGTGTTTCAAGCCAAGTAGATTG 1653

QY 502 aGluGluHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThrIleLysThrVally 522
Db 1654 TCGGATAAAGATACGTAACCTTCGAGACCAGTTGGATTCT---ACGGTTTGTACTATAAA 1710
```


LOCATION: (234187)..(234187)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (234220)..(234220)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (234814)..(234814)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (309398)..(309398)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (309418)..(309418)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (312837)..(312837)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (312993)..(312993)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (319226)..(319226)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (559167)..(559167)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (559241)..(559241)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (600992)..(600992)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (622708)..(622708)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (657081)..(657081)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (657203)..(657203)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (674435)..(674435)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (682442)..(682442)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (713652)..(713652)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (741684)..(741684)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (779455)..(779455)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (779676)..(779676)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (855539)..(855539)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (871619)..(871619)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1084830)..(1084830)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1096846)..(1096846)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1119881)..(1119881)

OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1130881)..(1130881)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1310988)..(1310988)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1313224)..(1313224)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1349473)..(1349473)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1349491)..(1349491)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1470091)..(1470091)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1569020)..(1569020)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1602912)..(1602912)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1603734)..(1603734)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1637998)..(1637998)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1664854)..(1664855)
OTHER INFORMATION: n equals a, t, c, or g
US-08-916-421B-1

Alignment Scores:

Pred. No.:	2.6e-29	Length:	1664976
Score:	392.50	Matches:	169
Percent Similarity:	40.05%	Conservative:	125
Best Local Similarity:	23.02%	Mismatches:	241
Query Match:	7.39%	Indels:	199
DB:	4	Gaps:	27

US-09-515-363C-2 (1-1025) x US-08-916-421B-1 (1-1664976)

Qy	306	LeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaLeuGluGlyLysAsnIle	325
Db	1658788	TTGGAGGCGAGGTTGTATCAGCAGATTATTGCAGCAATGCTTTAAAGAAAAAGACATTA	1658847
Qy	326	IleIleCysLeuProThrGlySerGlyLysThrArgValAlaValTyrIleAlaLysAsp	345
Db	1658848	TGTGTT---TTATCGACAGGTTTAGGTAAACAGCTATTGCTATTATTAGTTATAGCAGGT	1658904
Qy	346	HisLeuAspLysLysLysLysAlaSerGluProGlyLysValIleValLeuValAsnLys	365
Db	1658905	ATTTTAACAAAAAAG-----GATGGAAGGTTTAAATCTTAGCCCTTCA	1658949
Qy	366	ValLeuLeuValGluGlnLeuPheArgLysGluPheGlnPropheLeuLysLysTrpTyr	385
Db	1658950	AGACCTTTGGTTGAGCAACACTACAACAGATTAAACAGGTTTAAACATTGATGAAGAT	1659009
Qy	386	ArgValIleGlyLeuSerGlyAspThrGlnLeuLysIleSerPheProGluValValLys	405
Db	1659010	AAAATAATAGCTTTAACTGGAAAAATCCAGCCAAA---AAGAGAGCTGAACCTCTATAA	1659066
Qy	406	SerCysAspIleIleIleSerThrAlaGlnIleLeuGluAsnSerLeuLeuAsnGlu	425
Db	1659067	AAAGGGAAATCTTTATAGCTACACCACAGTTATAGAAAAACGATATCATA-----	1659117
Qy	426	AsnGlyGluAspAlaGly---ValGlnLeuSerAspPheSerLeuIleIleAspGlu	444
Db	1659118	-----GCTGGAAGATAAATGTGGATGAATTATTATTATTATGATAGCTGATGAA	1659165

QY 445 CysHisHisThrAsnLysGluAlaValTyrAsnAsnIleMetArgHisTyrLeuMetGln 464
Db 1659166 GCCCACCACACACAGGAGACCATGCCCTAT-----GCATTTGTAGCAAAA 1659210
QY 465 LysLeuLysAsnAsnArgLeuLysLysGluAsnLysProValIleProLeuProGlnIle 484
Db 1659211 AAATTTAAAGATAAA-----TGTCATATT 1659234
QY 485 LeuGlyLeuThrAlaSerProGlyValGlyGlyAlaThrLysGlnAlaLysAlaGluGlu 504
Db 1659235 TTAGGTTTAAACGGCATCTCCAGT-----TCTGATATTGAT 1659270
QY 505 HisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThrIleLysThrValLysGluAsn 524
Db 1659271 AAAGTCATGGAAATTTGTGAAACCTTAGCAATTGAGCACGTTGAAGTGAGAACTGAAGAT 1659330
QY 525 LeuAspGlnLeuLys----- 529
Db 1659331 GATGAGGATGTAAACCATACATCTGCTAAAGTAAACCTTATCCCAATTAGAAATTGATT 1659390
QY 530 ---AsnGlnIleGlnGluProCysLysLysPheAlaIleAlaAspAlaThrArgGlu--- 547
Db 1659391 CCCAACGAATTTAAAGAGCGTTAAAA-----TTAATAATGAAGCTTTAAAGGAGAGA 1659444
QY 548 -----AspPropheLysGluLys 553
Db 1659445 TTAAAAATATTAAAGATGCTGGAGTTATAAATTCATTGCCGATGTAAACAAAAACAGAA 1659504
QY 554 LeuLeuGluIleMetThrArgIleGlnThrTyrCysGlnMetSerProMetSerAspPhe 573
Db 1659505 CTTATTGAGCTAAATAATAAGCTATTTCCTAT----- 1659537
QY 574 GlyThrGlnProTyrGluGlnTrpAlaIleGlnMetGluLysLysAlaAlaLysLysGly 593
Db 1659537 ----- 1659537
QY 594 AsnArgLysGluArgValCysAlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIle 613
Db 1659538 -----GATGAAGAAGTGAAGTATGAACTTATAAAAGTTTGTTCAGAGGCTTTAAACTT 1659591
QY 614 AsnAspThrIleArgMetIleAspAla-----TyrThrHisLeuGlu 627
Db 1659592 ATGCATGCCAAAGAACTCTTAGAGAGCTCAAGGAAGAGTGTATTTTAAACTATATAAAT 1659651
QY 628 ThrPheTyrAsnGluGluLysAspLysLysPheAlaValIleGluAspAspSerAspGlu 647
Db 1659652 AAATTATCCATGCAAGAACAAATCAGCTAAATCTATTGTTAATGATGAA----- 1659702
QY 648 GlyGlyAspAspGluTyrCysAspGlyAspGluAspGluAspLeuLysLysProLeu 667
Db 1659702 ----- 1659702
QY 668 LysLeuAspGluThrAspArgPheLeuMetThrLeuPhePheGluAsnAsnLysMetLeu 687
Db 1659703 AAAGTTAGAGAGGCAGTTAATTTTATTAAATGAATCAGATGTAGAACATCCAAATTA--- 1659759
QY 688 LysArgLeuAlaGluAsnProGluTyrGluAsnGluLysLeuThrLysLeuArgAsnThr 707
Db 1659760 -----GGTAAAGTTGTTGATATGGTTAAAAAT 1659786
QY 708 IleMetGluGlnTyrThrArgThrGluGluSerAlaArgGlyIleIlePheThrLysThr 727
Db 1659787 ATTTTGGAAAAA-----AATAAGGATGAGAGAATTATTATCTTTGCTCAATAT 1659834
QY 728 ArgGlnSerAlaTyrAlaLeuSerGlnTrpIleThrGluAsnGluLysPheAlaGluVal 747
Db 1659835 AGGGACACTGTAGAGAAGATTGTTAATCTCTTAACTCAAAAT----- 1659876
QY 748 GlyValLysAlaHisHisLeuIleGlyAlaGlyHisSerSerGluPheLysProMetThr 767
Db 1659877 GGAATTAAGCAATAAGATTTATAGGACAGGCA---AATAAAGAGGAAGGGAATGAGT 1659933

QY 768 GlnAsnGluGlnLysGluValIleSerLysPheArgThrGlyLysIleAsnLeuLeuIle 787
Db 1659934 CAGAAAGAGCAAAATAGAAGCTATAGAGAGATTAAAAAA---GAGGGAAGTGTTTAGTT 1659990
QY 788 AlaThrThrValAlaGluGluGlyLeuAspIleLysGluCysAsnIleValIleArgTyr 807
Db 1659991 TCAACAAGCGTTTCTGAGGAGGAATAGATATTCCTCGGTAAATTACATCATATTTAT 1660050
QY 808 GlyLeuValThrAsnGluIleAlaMetValGlnAlaArgGlyArgAlaAspGlu 827
Db 1660051 GAACAGTGCCATCAGAAATTAGGTTTATTTCAGAGAGAGGTAGCGGTAGGGGAGAA 1660110
QY 828 -----SerThrTyrValLeuVal-----AlaHisSerGlySer 838
Db 1660111 GGAGGGAAGGTTTATGTTTAAATAGCTAAGGGAACAGCTGATGAAGCTTATTACAGGAGT 1660170
QY 839 GlyValIleGluHisGluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIle 858
Db 1660171 GCTTTATACAAAGAAAGGAGATGAAGAGATTATTAAAAAATATGTGTTATTGCTAAAT 1660230
QY 859 HisCysValGlnAsnMetLysProGluTyrAlaHisLysIleLeuGluLeuGlnMet 878
Db 1660231 AAGAGGTTACAGAAG---AAATTTGAAGAAAAATCTAAAGAGGAAATAAAGGAAGACA 1660287
QY 879 GlnSerIleMetGluLysLysMetLysThrLysArgAsnIleAlaLysHisTyrLys--- 897
Db 1660288 GAAGAAATAAAGAAAAAGAAATTTGAATCAAAAACTGCAGTAAAGAAACTAAGGAG 1660347
QY 898 -----AsnAsnProSerLeuIleThrPheLeuCysLysAsnCysSerVal 912
Db 1660348 GAAGAAGAAAAAACCAAAAGCCAGTAACGATATTAGATTTTCATTAAACAGATTGAAGTT 1660407
QY 913 -----LeuAlaCysSerGlyGluAsp----- 919
Db 1660408 AAGGAAAGGTCCTAAATCAGAAGAAGATAAAATAAAACAAGAGATAAAAAATCCGAAAAAG 1660467
QY 920 ---IleHisValIleGluLysMetHisHisValAsnMetThrPro----- 933
Db 1660468 CCAATAAAGATTATTGTAGATGTTAGAGAGAAGAATATGGCTAAGCTTTTACATAATTAT 1660527
QY 934 -----GluPheLysGluLeu-----TyrIleValArgGluAsnLysAla 946
Db 1660528 GCAATATTGAGCTAAAAACATTAGAAGTGGGAGATTATGTTTAAAGTATAGGTTAGTT 1660587
QY 947 LeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluIle 960
Db 1660588 GTTGAGAGAAAGACAGCTGAAGACTTTGTAAATTCAATAATT 1660629

RESULT 7
US-09-692-570-1
; Sequence 1, Application US/09692570
; Patent No. 6797466
; GENERAL INFORMATION:
; APPLICANT: Bult et al.
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanoco
; Patent No. 6797466
; TITLE OF INVENTION: jannaschii
; FILE REFERENCE: PB275C1
; CURRENT APPLICATION NUMBER: US/09/692,570
; CURRENT FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: US 60/024,428
; PRIOR FILING DATE: 1996-08-22
; PRIOR APPLICATION NUMBER: US 08/916,421
; PRIOR FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1664976
; TYPE: DNA
; ORGANISM: Methanococcus jannaschii
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (28222)..(28222)

```

; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (28257)..(28258)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (84773)..(84773)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (84808)..(84808)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (84812)..(84812)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98120)..(98120)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98159)..(98159)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98239)..(98239)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98266)..(98266)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98343)..(98343)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (103998)..(103998)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (148948)..(148948)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (163385)..(163385)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (191989)..(191989)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (191995)..(191995)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (231980)..(231980)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (234187)..(234187)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (234220)..(234220)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (234814)..(234814)
; OTHER INFORMATION: n equals a, t, c, or g

```

```

; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (309398)..(309398)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (309418)..(309418)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (312837)..(312837)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (312993)..(312993)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (319226)..(319226)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (559167)..(559167)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (559241)..(559241)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (600992)..(600992)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (622708)..(622708)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (657081)..(657081)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (657203)..(657203)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (674435)..(674435)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (682442)..(682442)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (713652)..(713652)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (741684)..(741684)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (779455)..(779455)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (779676)..(779676)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (855539)..(855539)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:

```


QY 788 AlaThrThrValAlaGluGluGlyLeuAspIleLysGluCysAsnIleValIleArgTyr 807
Db 1659991 TCAACAAGCGTTCTTGAGGAGGGAATAGATATTCCATCGGTAAATATACATCATATTTTAT 1660050
QY 808 GlyLeuValThrAsnGluIleAlaMetValGlnAlaArgGlyArgAlaAArgAlaAspGlu 827
Db 1660051 GAACCAAGTGCATCAGAAATTAGGTTTATTCAGAGGAGAGGTAGAGCGATGAGGGGAGAA 1660110
QY 828 -----SerThrTyrValLeuVal-----AlaHisSerGlySer 838
Db 1660111 GGAGGGAAGGTTTATGTTTAAATAGCTAAGGGAACAGCTGATGAAGCTTATTACAGGAGT 1660170
QY 839 GlyValIleGluHisGluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIle 858
Db 1660171 GCTTTATACAAGAAAGGAGATGAAGCAGATTATTAATAAATATGTGTTATTTGCTAAAT 1660230
QY 859 HisCysValGlnAsnMetLysProGluGluTyrAlaHisLysIleLeuGluLeuGlnMet 878
Db 1660231 AAGAGGTTACAGAAG---AAATTTGAAGAAATACTAAAGAGGAAATAAAGGAAGAGACA 1660287
QY 879 GlnSerIleMetGluLysLysMetLysThrLysArgAsnIleAlaLysHisTyrLys--- 897
Db 1660288 GAAGAAATAAAGAAAGAAATTTGAATCAAAAACTGCAGTAAAGAAAGAACTAAGGAG 1660347
QY 898 -----AsnAsnProSerLeuIleThrPheLeuCysLysAsnCysSerVal 912
Db 1660348 GAAGAAGAAAAACCAAAAGCCAGTAACGATATTAGATTTTCATTAAACAGATTGAAGTT 1660407
QY 913 -----LeuAlaCysSerGlyGluAsp----- 919
Db 1660408 AAGGAAAGGCTCTAAATCAGAAGAAGATAAAATAAACAGAGATAAAATTCGGAAGAAAG 1660467
QY 920 ---IleHisValIleGluLysMetHisHisValAsnMetThrPro----- 933
Db 1660468 CCAATAAAGATTATTGTAGATGTTAGAGAGAAGAAATATGGCTAAGCTTTTACATAATTAT 1660527
QY 934 -----GluPheLysGluLeu-----TyrIleValArgGluAsnLysAla 946
Db 1660528 GCAAAATATTGAGCTAAAAACATTAGAAAGTGGGAGATTATGTTTAAAGTGATAGGGTAGTT 1660587
QY 947 LeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluIle 960
Db 1660588 GTTGAGAGAAAGACAGCTGAAGACTTTCTGTAATTTCAATAATT 1660629

RESULT 8
US-09-853-768-3
; Sequence 3, Application US/09853768
; Patent No. 644466
; GENERAL INFORMATION:
; APPLICANT: Donna T. Ward
; APPLICANT: Andrew T. Watt
; TITLE OF INVENTION: ANTISENSE MODULATION OF HELICASE-MOI EXPRESSION
; FILE REFERENCE: RTS-0217
; CURRENT APPLICATION NUMBER: US/09/853,768
; CURRENT FILING DATE: 2001-05-10
; NUMBER OF SEQ ID NOS: 91
; SEQ ID NO 3
; LENGTH: 7037
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (183)....(5957)
US-09-853-768-3

Alignment Scores:
Pred. No.: 5,66e-27 Length: 7037
Score: 335.50 Matches: 186
Percent Similarity: 36.46% Conservative: 113
Best Local Similarity: 22.68% Mismatches: 233
Query Match: 6.32% Indels: 288
DB: 4 Gaps: 43

US-09-515-363C-2 (1-1025) x US-09-853-768-3 (1-7037)
QY 309 ArgProTyrGlnMetGluValAlaGlnProAlaLeuGluGlyLysAsnIleIleCys 328
Db 315 AGAAATATATCAGGTTGAACCTGTTGAAGCAGCTCTGGAT---CATAATACCATCGTCTGT 371
QY 329 LeuProThrGlySerGlyLysThrArgValAla---ValTyrIleAlaLysAsp----- 345
Db 372 TTAAACACTGGCTCAGGGAAGACATTATTGCTAGTACTACTCTACTAAAGAGCTGTCTC 431
QY 346 HisLeuAspLysLysLysLysAlaSerGluProGlyLys---ValIleValLeuValAsn 364
Db 432 TATCTAGATCTAGGGGAGACTTCAGCTAGAAAATGGAAAAAGGACGGTGTCTTCTGTCACAC 491
QY 365 LysVal---LeuLeuValGluGlnLeuPheArgLysGluPheGlnProPheLeuLysLys 383
Db 492 TCTGCAAAACCAGGTTGCTCAACAAGTGTCAAGTGTCAAGACTCATTCAAG--- 548
QY 384 TrpTyrArgValIleGlyLeuSerGlyAspThrGlnLeuLysIleSerPhe----- 400
Db 549 -----GTTGGGAATFACTCAAACTAGAACTAGAAATGCATCTTGGACAAAAGAG 596
QY 401 -----ProGluValValLysSerCysAspIleIleIleSerThrAlaGlnIleLeu 417
Db 597 AGATGGAACCAAGAGTTTACTAAGCAC---CAGGTTCTCATTTATGACTTGTCTATGTCGCC 653
QY 418 GluAsnSerLeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPhe 437
Db 654 TTGAATGTTTGAAAAAT-----GGTTACTTATCCTGTCAGACATT 695
QY 438 SerLeuIleIleIleAspGluCysHisHisThrAsnLysGluAlaValTyrAsnAsnIle 457
Db 696 AACCTTTTGGTGTGTTGATGAGTGTCTATCTTGCAATCTAGACCACCCCTATCGAGAAATTT 755
QY 458 MetArgHisTyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysLysGluAsnLysPro 477
Db 756 ATGAAG-----CTCTGTGAA----- 770
QY 478 ValIlePro---LeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyAla 496
Db 771 ATTTGTCCATCATGTCCTCGCATTTTGGGACTAAGTCTTCCATT---TTAAATGGGAAA 827
QY 497 ThrLysGlnAlaLysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeuAspAlaPhe 516
Db 828 TGGGATCCAGAGGATTTTGAAGAAAAAGTTTCAGAAACTA----- 866
QY 517 ThrIleLysThrValLysGluAsnLeuAspGln-----LeuLysAsn 530
Db 867 ---GAGAAAAATCTTAAAGAGTAATGCTGAAACTGCAACTGACCTGGTGGTCTTAGACAGG 923
QY 531 GlnIleGlnGluProCysLysLysPheAlaIleAlaAsp-----AlaThrArg 546
Db 924 TATACTTCTCAGCCATGTGAG-----ATTGTGGTGGATTGTGGACCATTCTTCTGACAGA 977
QY 547 GluAspProPheLysGluLysLeuLeuGluIle----- 557
Db 978 AGTGGGCTTTATGAAAGACTGCTGATGGAATTAGAAAGAGCAGCTTAATTTTATCAATGAT 1037
QY 558 -----MetThrArgIleGlnThrTyrCysGlnMetSerProMet 570
Db 1038 TGTAATATATCTGTACATTCAAAAAGAAAGAGATTCTACTTTAATTTTCGAAAACAGATACTA 1097
QY 571 SerAsp-----PheGlyThrGlnProTyrGluGlnTrpAlaIle 583
Db 1098 TCAGACTGTCGTGCCGTATTGTTAGTTCTGGGACCCCTGGTGTGCAGATAAAGTAGCTGGA 1157
QY 584 GlnMetGluLysLysAlaAlaLys-----LysGlyAsnArgLysGlu----- 597
Db 1158 ATGATGGTAAGAGAACTACAGAAATAACATCAACATGACGAAGAGGAGCTGCACAGGAAA 1217
QY 598 -----ArgValCysAlaGluHisLeu 604
Db 1218 TTTTATTGTTTACAGACACTTTCTCTAAGGAAAAATACATGCAGCTATGTGAGAGCACTTC 1277

QY 605 ArgLysTyrAsn-----GluAlaLeuGlnIleAsnAspThr 616
Db 1278 TCACCTGCCTCAGCTGAAATTTGTAACCTCTAAAGTAATCAAACTGCTCGAAATC 1337
QY 617 IleArgMetIleAspAlaTyrThrHis-----LeuGluThrPheTyrAsn 631
Db 1338 TTACGCAAAATATAAACCATATAGCGGACACAGTCTTGAAGCGTTGAGTGGTATATAAAT 1397
QY 632 GluGluLysAspLysPheAlaValIleGluAspAspSerAspGluGlyGlyAspAsp 651
Db 1398 AGAAATCAGGATAATTATGTGTATGG-----AGTGATTCTGAGGATGATGAT 1445
QY 652 GluTyrCysAspGlyAspGluAspGluAspLeuLysLysProLeuLysLeuAspGlu 671
Db 1446 GAG-----GATGAAGAAATTGAAGAAAAAGAGAGCCA----- 1478
QY 672 ThrAspArgPheLeuMetThrLeuPhePheGluAsnAsnLysMetLeuLysArgLeuAla 691
Db 1478 ----- 1478
QY 692 GluAsnProGluTyrGluAsnGluLysLeuThrLysLeuArgAsnThrIleMetGluGln 711
Db 1479 -----GAGACAAATTTTCCTCTCTCTTTTACCAACATTTTGTGC----- 1517
QY 712 TyrThrArgThrGluGluSerAlaArgGlyIleIlePheThrLysThrArgGlnSerAla 731
Db 1518 -----GGAATTATTTTGTGAAAGAAAGATACACAGCA 1550
QY 732 TyrAlaLeuSerGlnTrpIleThrGluAsnGluLys-----PheAla 745
Db 1551 GTTGTCTTAAACAGATTGATAAAGGAGCTGGCAACAGATCCAGAGCTGGCTTATATC 1610
QY 746 GluValGlyValLysAlaHisHisLeuIleGlyAlaGlyHis----- 759
Db 1611 AGTAGCAATTTTCATACTGGACATGGCATGGGAGGAGTGGCAACAGCTCGCAACACGATG 1670
QY 760 SerSerGluPheLysProMetThrGlnAsnGluGlnLysGluValIleSerLysPheArg 779
Db 1671 GAAGCAGAATTTCAGA-----AAACAGGAAGAGGTACTTAGGAATTCGA 1715
QY 780 ThrGlyLysIleAsnLeuIleAlaThrThrValAlaGluGluGlyLeuAspIleLys 799
Db 1716 GCACATGAGACCAACCTGCTTATTGCAACAAGTATTGTAGAAGAGGGTGTGTATATACCA 1775
QY 800 GluCysAsnIleValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAla 819
Db 1776 AAATGCAACCTGGTGGTTCGTTTGTGATTGCCACAGATATCGATCCTATGTTCAATCT 1835
QY 820 ArgGlyArgAlaArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGly 839
Db 1836 AAAGGAAGAGCAAGGGCACCCATCTCTAATATATATATATGTTAGCG----- 1880
QY 840 ValIleGluHisGluThrValAsnAspPheArgGluLysMet---MetTyrLysAlaIle 858
Db 1881 -----GATACAGACAAATAAAAAAGTTTGAAGAAGACCTTAAACCTACAAAGCTATT 1934
QY 859 HisCysValGlnAsnMetLysProGluGluTyrAlaHisLysIleLeuGluLeuGlnMet 878
Db 1934 ----- 1934
QY 879 GlnSerIleMetGluLysLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsn 898
Db 1935 -----GAAAAGATCTTGAGAAACAAG----- 1955
QY 899 AsnProSerLeuIleThrPheLeuCysLysAsnCysSerValLeuAlaCysSerGlyGlu 918
Db 1956 -----TGTTCCAAAGTCGGTTGATGACTGGTGAG 1982
QY 919 ---AspIleHisValIleGluLysMetHisHisValAsnMetThrProGluPheLysGlu 937
Db 1983 ACTGACATTGATCCTGTGATGGATGATGATCAGGTT---TTCCCACCA----- 2027

QY 938 LeuTyrIleValArgGluAsnLysAla----- 946
Db 2028 ---TATGTGTTGAGGCTGACCATGGTGGTCCAGGATGACAAATCAACACGGCCATTGGA 2084
QY 947 ---LeuGlnLysLysCysAla-AspTyrGlnIle-----As 957
Db 2085 CACATCAATAGATACTGTGCTAGATTACCAAGTGATCCGTTTACTCATCTAGTCTCTAAA 2144
QY 957 nGlyGlu-----IleIleCysLysCysG1 965
Db 2145 TGCAGAACCCGAGAGTTGCCTGATGGTACATTTTATTCAACTCTTTATCTGCCAATAAC 2204
QY 965 yGlnAlaTrp---GlyThrMetMetValHis-----LysG1 976
Db 2205 TCACCTCTTCGAGCCTCCATTGTTGGTCCACCAATGAGCTGTGTACGATTGGCTGAAAGA 2264
QY 976 yLeuAspLeuPro-----CysLeuLysIleArgAsnPheValValVal 990
Db 2265 GTTGTGCTCTCATTTGCTGTGAGAAACTGCACAAAATTTGGCGAACTGGATGACCATT 2322
RESULT 9
US-09-408-020-65
; Sequence 65, Application US/09408020
; Patent No. 6632937
; GENERAL INFORMATION:
; APPLICANT: Swanson, Ronald V.
; APPLICANT: Feldman, Robert A.
; APPLICANT: Schleper, Christa
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM
; FILE REFERENCE: DORP.002A
; CURRENT APPLICATION NUMBER: US/09/408,020
; CURRENT FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: 60/102,294
; PRIOR FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 65
; LENGTH: 1512
; TYPE: DNA
; ORGANISM: Cenarchaeum symbiosum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)... (1512)
US-09-408-020-65
Alignment Scores:
Pred. No.: 7.43e-23 Length: 1512
Score: 289.50 Matches: 137
Percent Similarity: 39.40% Conservative: 88
Best Local Similarity: 23.99% Mismatches: 188
Query Match: 5.45% Indels: 158
DB: 4 Gaps: 23
US-09-515-363C-2 (1-1025) x US-09-408-020-65 (1-1512)
QY 303 GluPro---GluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaLeuGlu 321
Db 34 GAGCCCGGCGCGCTCGAGAGCGCGACTACCGGTGGCCCTTCCGAGCAGGCCCATACGG 93
QY 322 GlyLysAsnIleIleIleCysLeuProThrGlySerGlyLysThrArgValAlaValTyr 341
Db 94 ---GAAAACCTGCATAGTGGTGTGCTACCGGCTCGGCAAGACGGCGGTGGCCCTGCAG 150
QY 342 IleAlaLysAspHisLeuAspLysLysLysAlaSerGluProGlyLysValIleVal 361
Db 151 GTGATCTCCCACTATTGGACGAAGGCGAGGGGGCTCTCTCTCTGCGCCGACAAGGGTG 210
QY 362 LeuValAsnLysValLeuLeuValGluGlnLeuPheArgLysGluPheGlnPropheLeu 381
Db 211 CTGGTAAAC-----CAGCACCGCCAGTTTCCTG 237
QY 382 LysLysTrpTyrArgValIleGlyLeuSerGlyAspThrGlnLeu-----LysIleSer 399

Db 238 GGC-----AGGGCCCTTACCATAATCCGATATTACCCCTGGTCACAGGGCAGGACACC 288

QY 400 PheProGluValLysSerCys-----AspIleIleIleSerThrAlaGlnIleLeu 417

Db 289 GTCCCGAGGGCGAAAAGCTTGGGGCGGCAGCGTGATCTGCGCCACCCCGAGATAACA 348

QY 418 GluAsnSerLeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPhe 437

Db 349 AGAAACGACATAGCGCGC-----GGATGTGTCCTCCGCTCGAACAGTTC 390

QY 438 SerLeuIleIleIleAspGluCysHisHisThrAsnLysGluAlaValTyrAsnAsnIle 457

Db 391 GGCCTGGTTGTTCGACGAGGCCACAGGGCGGTGGCGACTATGCTATTCGCAATA 450

QY 458 MetArgHisTyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysGluAsnLysPro 477

Db 451 GCG-----CGTGCACTGGGGGAGAACTCT--- 474

QY 478 ValIleProLeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyGlyAlaThr 497

Db 475 -----AGATGATCGGCATGCTGCG----- 495

QY 498 LysGlnAlaLysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThr 517

Db 496 -----ACC 498

QY 518 IleLysThrValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluProCysLys 537

Db 499 CTTCCAAGCGAGAGGGAGAAAGCCGACGAGATAATGGGCACTCTCTCTCAAAG----- 552

QY 538 LysPheAlaIleAlaAspAlaThrArgGluAspPro----- 549

Db 553 -----AGCATAGCACAAAGGACCCAGACGACCCGGATGTAAAGCCCTACGTGCAGGAG 606

QY 550 -----PheLysGluLysLeuLeuGluIleMetThrArgIleGlnThrTyr 564

Db 607 ACCGAAACTGAATGGATAAAGGTGGAGCTGCCCCCGGAGATGAAGGAGATCCAAAAGCTC 666

QY 565 CysGlnMetSerProMetSerAspPheGlyThrGlnProTyrGluGlnTrpAlaIleGln 584

Db 667 CTGAAGATG-----GCCCTCGAC 684

QY 585 MetGluLysLysAlaAlaLysLys-----GlyAsnArg-LysGluAr 598

Db 685 GAAAGATATCGGGCCCTCAAGAGGTGCGGCTATGATCTCGGCTCGAACAGGTGCTCTCG 744

QY 598 gValCysAlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleAr 618

Db 745 GCTCTGCTCCGCCCTTCGCATGGTCTTAAGCGGCAACAGCGGGCGGCAAGCCTTTG 804

QY 618 gMetIleAspAlaTyrThrHisLeuGluThrPheTyrAsnGluGluLys---AspLysLy 637

Db 805 TTTACTGCGATACGCATCACATACGCGCTCAACATATTCAGGCCCCACGGGTCAACCGG 864

QY 637 sphe-AlaValIleGluAspAspSerAspGluGlyGlyAspAspGluTyrCysAspGlyA 657

Db 865 TTTCTAAAGTTCTGCGAGAGGACCGTCAAGAAAAAGGG----- 902

QY 657 spGluAspGluAspLeuLysLys-ProLeuLysLeuAspGluThrAspArgPheLeu 676

Db 903 -----CGCCGGTGTTCAGAGCTGTTCGAGGAGGACAGA----- 936

QY 677 MetThrLeuPhePheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnProGluTyr 696

Db 937 -----AACTTTACAGGGGCCATGGCGCGCGCAAGCGCGCAGGCAGCCGGCATG 987

QY 697 GluAsnGluLysLeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArgThrGlu 716

Db 988 GAGCATCCAAAGATACCAAGTTG-----GAA 1014

QY 717 GluSerAlaArgGly-----IleIlePheThrLysThrArgGlnSerAla 731

Db 1015 GAGGCTGTGCGGGGGCCAAAGGGAAGGCGCTGGTCTTTACAAGCTACAGGACTCTGTC 1074

QY 732 TyrAlaLeuSerGlnTrpIleThrGluAsnGluLysPheAlaGluValGlyValLysAla 751

Db 1075 GATTTAATA-----CACTCAAAGCTGCAGGCTGCCGGGATAAACTCG 1116

QY 752 HisHisLeuIleGly---AlaGlyHisSerSerGluPheLysProMetThrGlnAsnGlu 770

Db 1117 GGGATCCTCATAGGAAGGCGGGAGAA-----AAGGGCCTCAAGCAGAAAAAA 1164

QY 771 GlnLysGluValIleSerLysPheArgThrGlyLysIleAsnLeuLeuIleAlaThrThr 790

Db 1165 CAGGTAGAGACTGTCCCAAGTTCGCGCAGCGGGGATACGACGTGCTCGTATCTACAAGA 1224

QY 791 ValAlaGluGluGlyLeuAspIleLysGluCysAsnIleValIleArgTyrGlyLeuVal 810

Db 1225 GTGGCGAGGAGGGCCTCGACATATCGAGGTAAACCTTGTGGTATTCTATGACAATGTC 1284

QY 811 ThrAsnGluIleAlaMetValGlnAlaArgGlyArg---AlaArgAlaAspGluSerThr 829

Db 1285 CCAAGCTCGATAAGGTATGTGCAGAGAAGGGCGAGGACCGGCAAGGACCGGGCAAG 1344

QY 830 TyrValLeuValAlaHisSerGlySer 838

Db 1345 CTGGTGTACTGATGGCAAAGGGGACT 1371

RESULT 10

US-09-408-020-1

; Sequence 1, Application US/09408020

; Patent No. 6632937

; GENERAL INFORMATION:

; APPLICANT: Swanson, Ronald V.

; APPLICANT: Feldman, Robert A.

; APPLICANT: Schleper, Christa

; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM

; FILE REFERENCE: DORP.002A

; CURRENT APPLICATION NUMBER: US/09/408,020

; CURRENT FILING DATE: 1999-09-29

; PRIOR APPLICATION NUMBER: 60/102,294

; PRIOR FILING DATE: 1998-09-29

; NUMBER OF SEQ ID NOS: 123

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 1

; LENGTH: 32998

; TYPE: DNA

; ORGANISM: Cenarchaeum symbiosum

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (7604)...(8908)

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (8961)...(9767)

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (10545)...(10922)

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (13944)...(14612)

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (18638)...(20149)

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (20554)...(20955)

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (20956)...(21834)

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (25151)...(26377)

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (27535)...(28002)

; FEATURE:

; NAME/KEY: CDS			
; LOCATION: (28065) ... (29483)			
US-09-408-020-1			
Alignment Scores:			
Pred. No.:	1.84e-20	Length:	32998
Score:	289.50	Matches:	137
Percent Similarity:	39.40%	Conservative:	88
Best Local Similarity:	23.99%	Mismatches:	188
Query Match:	5.45%	Indels:	158
DB:	4	Gaps:	23
US-09-515-363C-2 (1-1025) x US-09-408-020-1 (1-32998)			
QY	303	GluPro---GluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaLeuGlu	321
Db	18671	GAGCCCGCGCGTCGAGAGGCGCGACTACCAAGTGGCGCTTCCGAGCAGGCCATACGG	18730
QY	322	GlyLysAsnIleIleCysLeuProThrGlySerGlyLysThrArgValAlaValTyr	341
Db	18731	---GAAAACTGCATAGTGGTCTACCGCCTCGGCAAGACGGCCGTGGCCCTGCAG	18787
QY	342	IleAlaLysAspHisLeuAspLysLysLysAlaSerGluProGlyLysValIleVal	361
Db	18788	GTGATCTCCCACTATTGGACGAAGCAGCGGGGGCTCTCTCTCTTGGCCCGACAAGGGTG	18847
QY	362	LeuValAsnLysValLeuLeuValGluGlnLeuPheArgLysGluPheGlnProPheLeu	381
Db	18848	CTGGTAAAC-----CAGCACCGCCAGTTCCTG	18874
QY	382	LysLysTrpTyrArgValIleGlyLeuSerGlyAspThrGlnLeu-----LysIleSer	399
Db	18875	GGC-----AGGGCCCTTACCATATCCGATATTACCTGGTCACAGGCGAGGACACC	18925
QY	400	PheProGluValValLysSerCys-----AspIleIleIleSerThrAlaGlnIleLeu	417
Db	18926	GTCCCGAGGCGCAAAAGCTTGGGCGGCAGCGTGATCTGGCCACCCCGAGATAACA	18985
QY	418	GluAsnSerLeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPhe	437
Db	18986	AGAAACGACATAGCGCGC-----GGATGGTCCCGCTCGAACAGTTC	19027
QY	438	SerLeuIleIleIleAspGluCysHisHisThrAsnLysGluAlaValTyrAsnAsnIle	457
Db	19028	GGCCTGGTTGTTCGACGAGGCCACAGGGCGGTGGCGCACTATGCCTATTCCCAATA	19087
QY	458	MetArgHisTyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysLysGluAsnLysPro	477
Db	19088	GCG-----CGTGAGTGGGGGAGAACTCT---	19111
QY	478	ValIleProLeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyGlyAlaThr	497
Db	19112	-----AGAATGATCGGCATGACTGCG-----	19132
QY	498	LysGlnAlaLysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThr	517
Db	19133	-----ACC	19135
QY	518	IleLysThrValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluProCysLys	537
Db	19136	CTTCCAAGCGAGAGGGAGAAAGCCGACGAGATAATGGGCACCTTCTCTCTCAAAG-----	19189
QY	538	LysPheAlaIleAlaAspAlaThrArgGluAspPro-----	549
Db	19190	-----AGCATAGCACAAAGGACCGGACGACCGCGGATGTAAAGCCCTACGTGCAGGAG	19243
QY	550	-----PheLysGluLysLeuLeuGluIleMetThrArgIleGlnThrTyr	564
Db	19244	ACCGAAACTGAATGATAAAGGTGGAGCTGCCCCCGGAGATGAAGGAGATCCAAAAGCTC	19303
QY	565	CysGlnMetSerProMetSerAspPheGlyThrGlnProTyrGluGlnTrpAlaIleGln	584
Db	19304	CTGAAGATG-----GCCCTCGAC	19321
QY	585	MetGluLysLysAlaAlaLysLys-----GlyAsnArg-LysGluAr	598
Db	19322	GAAAGATATGCGCCCTCAAGAGGTGCGGTATGATCTCGGCTCGAACAGGTCTCTCG	19381
QY	598	gValCysAlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleAr	618
Db	19382	GCTCTGCTCCGCTTCGCATGGTCTTCTTAAGCGGCAACAGCGGCGGCAAAAGCCTTG	19441
QY	618	gMetIleAspAlaTyrThrHisLeuGluThrPheTyrAsnGluGluLys---AspLysLy	637
Db	19442	TTTACTCGGATACGCATCACATACCGCTCAACATATTTCGAGGCCACCGGGTCACGCCG	19501
QY	637	sPhe-AlaValIleGluAspAspSerAspGluGlyGlyAspAspGluTyrCysAspGlyA	657
Db	19502	TTTCTAAAGTTCTGCGAGAGGACCGTCAAGAAAAAGGG-----	19539
QY	657	spGluAspGluAspAspLeuLysLys-ProLeuLysLeuAspGluThrAspArgPheLeu	676
Db	19540	-----CGCCGGTGTTCGACAGCTGTTTCGAGGAGGACAGA-----	19573
QY	677	MetThrLeuPhePheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnProGluTyr	696
Db	19574	-----AACTTTACAGGGGCCATGGCGCGCAAGCGCGCAGGACCCGGCATG	19624
QY	697	GluAsnGluLysLeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArgThrGlu	716
Db	19625	GAGCATCCAAAGATACCAAAAGTTG-----GAA	19651
QY	717	GluSerAlaArgGly-----IleIlePheThrLysThrArgGlnSerAla	731
Db	19652	GAGGTGTGCGCGGGCCAAAGGAAGCGCTGTCTTTACAAGCTACAGGACTCTGTC	19711
QY	732	TyrAlaLeuSerGlnTrpIleThrGluAsnGluLysPheAlaGluValGlyValLysAla	751
Db	19712	GATTTAATA-----CACTCAAAGCTGCAGGCTGCCGGGATAAACTCG	19753
QY	752	HisHisLeuIleGly---AlaGlyHisSerSerGluPheLysProMetThrGlnAsnGlu	770
Db	19754	GGGATCTCTATAGGAAAGCGCGGAGAA-----AAGGGCTCAAACGAGAAAAA	19801
QY	771	GlnLysGluValIleSerLysPheArgThrGlyLysIleAsnLeuLeuIleAlaThrThr	790
Db	19802	CAGTAGAGACTGTCCCAAGTTCGCGACGCGGGGATACGACGTGCTGTATCTACAAGA	19861
QY	791	ValAlaGluGluGlyLeuAspIleLysGluCysAsnIleValIleArgTyrGlyLeuVal	810
Db	19862	GTGGCGGAGGAGGCGCTCGACATATCGGAGGTAAACCTTGTGGTATTCTATGACAATGTC	19921
QY	811	ThrAsnGluIleAlaMetValGlnAlaArgGlyVArg--AlaArgAlaAspGluSerThr	829
Db	19922	CCAAGTCGATAAGGTATGTGCAGAGAAGGGGCGAGGCCGCGCAGGAAGGACCGGGCAAG	19981
QY	830	TyrValLeuValAlaHisSerGlySer	838
Db	19982	CTGGTGGTACTGATGGCAAGGGGACT	20008
RESULT 11			
US-09-408-020-33			
; Sequence 33, Application US/09408020			
; Patent No. 6632937			
; GENERAL INFORMATION:			
; APPLICANT: Swanson, Ronald V.			
; APPLICANT: Feldman, Robert A.			
; APPLICANT: Schleper, Christa			
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM			
; FILE REFERENCE: DCORP.002A			
; CURRENT APPLICATION NUMBER: US/09/408,020			
; CURRENT FILING DATE: 1999-09-29			
; PRIOR APPLICATION NUMBER: 60/102,294			
; PRIOR FILING DATE: 1998-09-29			
; NUMBER OF SEQ ID NOS: 123			
; SOFTWARE: FastSeq for Windows Version 3.0			

```
; SEQ ID NO 33
; LENGTH: 1509
; TYPE: DNA
; ORGANISM: Cenarchaeum symbiosum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1509)
US-09-408-020-33

Alignment Scores:
Pred. No.:      3.2e-22      Length:      1509
Score:          284.00      Matches:      133
Percent Similarity: 41.42%   Conservative: 94
Best Local Similarity: 24.27% Mismatches:      207
Query Match:      5.35%     Indels:        114
DB:               4         Gaps:          23

US-09-515-363C-2 (1-1025) x US-09-408-020-33 (1-1509)

QY 303 GluPro---GluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaLeuGlu 321
D 34 GAGCCCGGTGCGTCGAGAGCGCGACTACCAAGTGGGCTGGCGGAACAGGCCATACGG 93
QY 322 GlyLysAsnIleIleCysLeuProThrGlySerGlyLysThrArgValAlaValTyr 341
D 94 ---GAGAACTGTATCGTGGTCTCCCGACGGGCTCGGCAAGACTGCCGTGCCCTCCAG 150
QY 342 IleAlaLysAspHisLeuAspLysLysLysLysAlaSerGluProGlyLysValIleVal 361
D 151 GTGATCGCCCACTATCTGACGAGGGCGCGGGCGCTTCTTCTTGGCCCTACAAGGGTC 210
QY 362 LeuValAsnLysValLeuLeuValGluGlnLeuPheArgLysGluPheGlnPropheLeu 381
D 211 CTGGTAAAC-----CAGCACCGCCAGTTCTCTG 237
QY 382 LysLysTyrTyrArgValIleGlyLeuSerGlyAspThrGlnLeu-----LysIleSer 399
D 238 GGC-----AGGGCCCTTACCATATCCGATATTACACTGGTCACGGGAGGACACC 288
QY 400 PheProGluValValLysSerCys-----AspIleIleSerThrAlaGlnIleLeu 417
D 289 ATTCCCGCGCGCAAAAGCGTGGGAGGAGCGGTGATCTGCGCCACGCCCGAGATAGCA 348
QY 418 GluAsnSerLeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPhe 437
D 349 AGAAATGATATA-----GAGCGGGCCTGGTCCCGCTCGAACAGTTC 390
QY 438 SerLeuIleIleIleAspGluCysHisHisThrAsnLysGluAlaValTyrAsnAsnIle 457
D 391 GGCCTGGTCATATTCGACGAGGCCACACAGGGCGGTGGCGACTATGCTTCTTCCATA 450
QY 458 MetArgHisTyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysLysGluAsnLysPro 477
D 451 GCGCGG-----GCG 459
QY 478 ValIleProLeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyAlaThr 497
D 460 GTAGGGGATAACTCCAGGATGGTGGGCATGACTGCGACGCTT-----CCCAGC 507
QY 498 LysGlnAlaLysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThr 517
D 508 GAGAGGGAGAGGACAGCAGAG-----ATAATGGGCACCTGCTCTCCAGGAGC 555
QY 518 IleLysThrValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluProCysLys 537
D 556 ATAGCCAGAGGACAGAAGACGACCCGGACGTAAGGCCCTATGTACAGGAGACTGCCACC 615
QY 538 LysPheAlaIleAlaAspAlaThrArgGluAspPropheLysGluLysLeuLeuGluIle 557
D 616 GACTGGATAAAGGTGGATCTTCCCCCGAGATGAAGAGATACAGAGGCTCCTCAAGCTG 675
QY 558 MetThrArgIleGlnThrTyrCysGlnMetSerProMetSer---AspPheGlyThrGln 576
D 558 MetThrArgIleGlnThrTyrCysGlnMetSerProMetSer---AspPheGlyThrGln 576
```

```
D 676 GCCCTCGAC---GAGAGGTATTCTCTCCCTCAAGAGGTGCGGGTACGATCTTGGCTCGAAC 732
QY 577 ProTyrGluGlnTrpAlaIleGlnMetGluLysLysAlaLysLysGlyAsnArgLys 596
D 733 AGGTGCTCTCGGCGCTGCTCCGGCTG---CGCATGGTGGTGTGGCGGCAACAGGCGC 789
QY 597 GluArgValCysAlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThr 616
D 790 GCG-----GCCAAGCGCTGTTTCACTGGCATACGCATAACGTACGCG 831
QY 617 IleArgMetIleAspAlaTyr-----ThrHisLeuGluThrPheTyrAsnGluGluLys 634
D 832 CTAACATATTCGAGGCGCACGGGGTCAAGCCCTTTCTAAAGTTCTCGAGAGGACCTCC 891
QY 635 AspLysLysPheAlaValIleGluAspAspSerAspGluGlyGlyAspAspGluTyrCys 654
D 892 AAGAAAAAGGCGTGGCGTGGCGGAG----- 918
QY 655 AspGlyAspGluAspGluAspAspLeuLysLysProLeuLysLeuAspGluThrAspArg 674
D 919 -----CTGTTTCAACAGGACCGG 936
QY 675 PheLeuMetThrLeuPhePheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnPro 694
D 937 -----AACTTTACAGGGGCCATCGCGCGCGCAAGGCGCGCAGCGGCA 981
QY 695 GluTyrGluAsnGluLysLeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArg 714
D 982 GGCATGGAGCATCCCAAGATACCAAGCTCGAGGATGCCGTC----- 1023
QY 715 ThrGluGluSerAlaArgGly-----IleIlePheThrLysThrArgGlnSerAlaTyr 732
D 1024 -----CGCGGGCGCGGAAAGCGCTGGTCTTTACGAGCTATCGTGATCTGTGCGAC 1077
QY 733 AlaLeuSerGlnTrpIleThrGluAsnGluLysPheAlaGluValGlyValLysAlaHis 752
D 1078 CTCATA-----CACTCAAGACTCAAGCGCGCGGGATAAACTCGGGC 1119
QY 753 HisLeuIleGly---AlaGlyHisSerSerGluPheLysProMetThrGlnAsnGluGln 771
D 1120 ATCCTGATAGGAAAGCGGGAGAA-----AAGGGCTAAAGCAGAGAAACAG 1167
QY 772 LysGluValIleSerLysPheArgThrGlyLysIleAsnLeuLeuIleAlaThrThrVal 791
D 1168 GTGAGACTGTGGCAAAGTTCCGTGACGGCGGGTACGACGTGCTGGTATCGACGAGGTC 1227
QY 792 AlaGluGluGlyLeuAspIleLysGluCysAsnIleValIleArgTyrGlyLeuValThr 811
D 1228 GCGAGGAGGGGCTCGACATATCGAGGTCAACCTGGTGATATTTCTATGACAATGTGCCA 1287
QY 812 AsnGluIleAlaMetValGlnAlaArgGlyArg---AlaArgAlaAspGluSerThrTyr 830
D 1288 AGCTCGATCAGGTACGTGCAGAGAGGGGGAGAACAGGCAAGAGACGCCCGGCGGCTG 1347
QY 831 ValLeuValAlaHisSerGlySer 838
D 1348 ATAGTATTGATGGCAAAGGGGACG 1371
```

RESULT 12

```
US-09-408-020-2
; Sequence 2, Application US/09408020
; Patent No. 6632937
; GENERAL INFORMATION:
; APPLICANT: Swanson, Ronald V.
; APPLICANT: Feldman, Robert A.
; APPLICANT: Schleper, Christa
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM
; FILE REFERENCE: DCOMP.002A
; CURRENT APPLICATION NUMBER: US/09/408,020
; CURRENT FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: 60/102,294
; PRIOR FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 123
```


Db 35678 ATCTGATAGGAAAGCGGGAGAA-----AAGGGCCTAAACGAGAAACAG 35725
Qy 772 LysGluValIleSerLysPheArgThrGlyLysIleAsnLeuLeuIleAlaThrThrVal 791
Db 35726 GTGGAGACTGTGGCAAGTCCGTGACGGCGGTACGACGTGCTGGTATCGACGAGGTC 35785
Qy 792 AlaGluGluGlyLeuAspIleLysGluCysAsnIleValIleArgTyrGlyLeuValThr 811
Db 35786 GCGGAGGAGGGCTCGACATATCGGAGGTCAACCTGGTGATATTCTATGACAAATGTGCCA 35845
Qy 812 AsnGluIleAlaMetValGlnAlaArgGlyArg---AlaArgAlaAspGluSerThrTyr 830
Db 35846 AGTCGATCAGGTACGTGCAGAGGAGGGGGAGAACAGGCAGACAGGACCGCGCAGGCTG 35905
Qy 831 ValLeuValAlaHisSerGlySer 838
Db 35906 ATAGTATTGATGGCAAGGGGACG 35929

RESULT 13
US-08-143-576-6
; Sequence 6, Application US/08143576
; Patent No. 5643761
; GENERAL INFORMATION:
; APPLICANT: Fisher, Paul B.
; APPLICANT: Jiang, Hongping
; TITLE OF INVENTION: METHOD FOR GENERATING A SUBTRACTED
; TITLE OF INVENTION: CDNA LIBRARY AND USES OF THE GENERATED LIBRARY
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John P. White, c/o Cooper & Dunham
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10112

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/143,576
; FILING DATE: 25-OCT-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 43563/JPW/AKC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 977-9550
; TELEFAX: (212) 664-0525
; TELEX: 422523 COOP UI
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 193 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO

Alignment Scores:
Pred. No.: 2,71e-17 Length: 193
Score: 227.50 Matches: 50
Percent Similarity: 75.36% Conservative: 2
Best Local Similarity: 72.46% Mismatches: 10
Query Match: 4.28% Indels: 8
DB: 1 Gaps: 1

US-09-515-363C-2 (1-1025) x US-08-143-576-6 (1-193)

Qy 947 LeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluIleIleCysLysCysGlyGln 966
Db 1 CTGCAAAAGAGTGTGCCGACTATAA-ATAAATGGTGAATCATCTGCAAAATGTGGCCAG 59
Qy 967 AlaTrpGlyThrMetMetValHisLysGlyLeuAspLeuProCysLeuLysIleArgAsn 986
Db 60 GCTTGGGGAACAATGATGGTGCACAAAGGCTTAGATTTGCCTTGTCTCAAAATAAGGAAT 119
Qy 987 PheValValPheLysAsnAsnSerThrLysLysGlnTyrLysLysTrpValGluLeu 1006
Db 120 TTTGTAGTGGTTCAAAA-----TATCACAAGAACGTACAAAGTG 158
Qy 1007 ProIleThrPheProAsnLeuAspTyr 1015
Db 159 GTAGATACTATCACATTCACTGACTAT 185

RESULT 14
US-09-221-268D-11
; Sequence 11, Application US/09221268D
; Patent No. 6720408
; GENERAL INFORMATION:
; APPLICANT: Fisher, Paul B.
; APPLICANT: Jiang, Hongping
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES WHICH SPECIFICALLY HYBRIDIZE WITH
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCE ENCODING MDA-7 AND PHARMACEUTICAL COMPOSI
; FILE REFERENCE: A34534-A-A (070050.1637)
; CURRENT APPLICATION NUMBER: US/09/221,268D
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: 08/316,537
; PRIOR FILING DATE: 1994-09-30
; PRIOR APPLICATION NUMBER: 08/143,576
; PRIOR FILING DATE: 1993-10-27
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 193
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-221-268D-11

Alignment Scores:
Pred. No.: 2,71e-17 Length: 193
Score: 227.50 Matches: 50
Percent Similarity: 75.36% Conservative: 2
Best Local Similarity: 72.46% Mismatches: 10
Query Match: 4.28% Indels: 8
DB: 4 Gaps: 1

US-09-515-363C-2 (1-1025) x US-09-221-268D-11 (1-193)

Qy 947 LeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluIleIleCysLysCysGlyGln 966
Db 1 CTGCAAAAGAGTGTGCCGACTATAA-ATAAATGGTGAATCATCTGCAAAATGTGGCCAG 59
Qy 967 AlaTrpGlyThrMetMetValHisLysGlyLeuAspLeuProCysLeuLysIleArgAsn 986
Db 60 GCTTGGGGAACAATGATGGTGCACAAAGGCTTAGATTTGCCTTGTCTCAAAATAAGGAAT 119
Qy 987 PheValValPheLysAsnAsnSerThrLysLysGlnTyrLysLysTrpValGluLeu 1006
Db 120 TTTGTAGTGGTTCAAAA-----TATCACAAGAACGTACAAAGTG 158
Qy 1007 ProIleThrPheProAsnLeuAspTyr 1015
Db 159 GTAGATACTATCACATTCACTGACTAT 185

RESULT 15
US-09-595-684B-30
; Sequence 30, Application US/09595684B
; Patent No. 6544766
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe

```

; APPLICANT: Ohashi, Cara
; APPLICANT: Sakowicz, Roman
; APPLICANT: Vaisberg, Eugeni
; APPLICANT: Wood, Kenneth
; APPLICANT: Yu, Ming
; TITLE OF INVENTION: Human kinesins and methods of producing
; TITLE OF INVENTION: and purifying human kinesins
; FILE REFERENCE: cytop036
; CURRENT APPLICATION NUMBER: US/09/595,684B
; CURRENT FILING DATE: 2002-06-24
; PRIOR APPLICATION NUMBER: 09/295,612
; PRIOR FILING DATE: 2000-04-20
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 8257
; TYPE: DNA
; ORGANISM: Human
US-09-595-684B-30

Alignment Scores:
Pred. No.: 3,19e-13 Length: 8257
Score: 217.50 Matches: 209
Percent Similarity: 35.12% Conservative: 185
Best Local Similarity: 18.63% Mismatches: 421
Query Match: 4.10% Indels: 307
DB: 4 Gaps: 47

US-09-515-363C-2 (1-1025) x US-09-595-684B-30 (1-8257)
QY 9 GluAsnPheArgTyrLeuIleSerCysPheArgAlaArgValLysMetTyrIleGlnVal 28
Db 3208 GAGCAACAAGGAAGATATTTCTTTAAATACAGAGAGAAAATGAACCTCCAACAATGTTA 3267
QY 29 GluProValLeuAspTyrLeuThrPheLeuProAlaGluValLysGluGlnIleGlnArg 48
Db 3268 GAGAGTGTATAGCAGAAAAGGAACAATTGAAGACTGACCTAAAGGAAAATATTGAAATG 3327
QY 49 ThrValAlaThrSerGlyAsnMetGlnAlaValGluLeuLeuLeuSerThrLeuGluLys 68
Db 3328 ACCATTGAA-----AACCAGGAAGAATTAAAGACTTCTTGGGGATGAACCTTAAAAAG 3378
QY 69 GlyValTrpHisLeuGlyTrpThrArgGluPheVal-----GluAlaLeu 83
Db 3379 -----CAACAAGAGATAGTTGCACAAGAAAAGAACCATGCCATA 3417
QY 84 ArgArgThrGlySerProLeuAlaAlaArgTyrMetAsnProGluLeuThrAspLeuPro 103
Db 3418 AAGAAAGAAGGAGAGCTTTCTAGGACCTGTGACAGACTGGCAGAGAGTTGAAGAAAATA 3477
QY 104 SerProSerPheGluAsnAlaHisAspGluTyrLeuGlnLeuLeuAsnLeuLeuGlnPro 123
Db 3478 AAGGAAAAGAGCCAGCAACTCCAAGAAAACAGCAACAACCTTCTTAATGTACAAGAGAG 3537
QY 124 ThrLeuValAspLysLeuLeuValArgAspVal-----LeuAspLysCysMetGluGlu 141
Db 3538 ATGAGTGAGATGCAGAAAAGATTAATGAATAGAGAAATTAAAGAAATGAATTAAGAAGAC 3597
QY 142 GluLeuLeuThrIleGlu-----AspArgAsnArgIleAlaAlaAlaGluAsn---Asn 158
Db 3598 AAAGAATTGACATTGGAAACATATGGAACACAGAGAGGCTTGAGTTGGCTCAGAAAACCTTAAT 3657
QY 159 GlyAsnGluSerGlyValArgGluLeuLys---ArgIleValGlnLysGlu----- 175
Db 3658 GAAAAATTATGAGGAAGTGAATCTATAACCAAGAAAAGAAAGTTCTTAAGGAATTACAG 3717
QY 176 -----AsnTrpPheSerAlaPheLeuAsnValLeuArgGlnThr 188
Db 3718 AAGTCATTTGAAACAGAGAGAGACCCTTTAGAGGATATATAAGAGAAATTTGAAGCTACA 3777
QY 189 Gly-----AsnAsnGlu 192
Db 3778 GGCCTACAAACCAAGAAAGAACTAAAAAATTGCTCATATTCCCTTAAAGAAACCAAGAA 3837

```

```

QY 193 LeuValGlnGluThrGlySerAspCysSerGluSerAsnAlaGluIleGluAsnLeu 212
Db 3838 ACTATTGATGAAGAAAGAGC---GTATCTGAGAAGACAGCTCAATAATAATAACT 3894
QY 213 SerGlnValAspGlyProGlnVal-----Glu 221
Db 3895 CAGGACTTAGAAAAATCCCATACCAAAATTACAAGAAGAGATCCCAGTGCTTCATGAGGAA 3954
QY 222 GluGlnLeuLeuSerThrThrValGlnProAsnLeuGluLysGluValTrpGlyMetGlu 241
Db 3955 CAAGAGTTACTG-----CCTAATGTGAAAAAAGTCAGTGAAGACTCAGGAA 3999
QY 242 AsnAsnSerSerGluSerSerPheAlaAspSerSerValValSerGluSerAspThrSer 261
Db 4000 ACAATGAATGAACCTGGAGTTATTAAACAGAACAGTCCACAACCAAGGACTCA---ACAACA 4056
QY 262 LeuAlaGluGlySerValSerCysLeuAspGluSerLeuGlyHisAsnSerAsnMetGly 281
Db 4057 CTGGCAAGAATAGAAATG-----GAAAGGCTCAGGTTGAATGAAAAATTT--- 4101
QY 282 SerAspSerGlyThrMetGlySerAspSerAspGluGluAsnValAlaAlaArgAlaSer 301
Db 4102 -----CAAGAAAAGTCAGGAAGAGAGATAAAATCTCTAACCAAG 4137
QY 302 ProGluProGluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaLeuGlu 321
Db 4138 GAAAGAGACAACCTTTAAACAGATAAAAGAGCCCTTGAAGTTAAACATGACCAAGCTG--- 4194
QY 322 GlyLysAsnIleIleIleCysLeuProThrGlySerGlyLysThrArgValAlaValTyr 341
Db 4195 -----AAAGAACATATTAGAGAACT 4215
QY 342 IleAlaLys-----AspHisLeuAspLysLysLys 352
Db 4216 TTGGCTAAAATCCAGGAGTCTCAAGCAAAACAAGAACAGTCTCTTAATATGAAGAAAAA 4275
QY 353 AlaSerGluProGlyLysValIleValLeuValAsnLysValLeuValGluGlnLeu 372
Db 4276 GACAATGAAACTACCAAAATC-----GTGAGTGAGATGGAGCAA--- 4314
QY 373 PheArgLysGluPheGlnProPheLeuLysLysTrpTyrArgValIleGlyLeuSerGly 392
Db 4315 TTCAAAACCCAAAGATTTCAGCACTACTAAGGATAGAAATAGAAATGCTCGGATGTCC--- 4371
QY 393 AspThrGlnLeuLysIleSerPheProGlu-----ValLysSerCysAspIle 409
Db 4372 ---AAAAGACTTCAAGAAAGTCATGATGAAATGAAATCTGTAGCTAAGGAGAAAGATGAC 4428
QY 410 IleIleSerThrAlaGlnIleLeuGluAsnSerLeuLeuAsnLeuGluAenGly----- 427
Db 4429 CTACAGAGCTGCAAGAAGTCTTCAATCTGAAAGTGACCAGCTCAAAGAAAACATAAAA 4488
QY 428 GluAspAlaGlyValGlnLeuSerAspPheSerLeuIleIleIleAspGluCysHisHis 447
Db 4489 GAAATTGTAGCTAAACACCTGGAAACTGAAGAGGAACTTAAAGTTGCTCATTTGTCCTG 4548
QY 448 ThrAsnLysGluAlaValTyrAsnAsnIleMetArgHisTyrLeuMetGlnLysLeuLys 467
Db 4549 AAAGAACAGAGGAAACTATTAAAT-----GAGTTAAGA 4581
QY 468 AsnAsnArgLeuLysLysGluAsnLysProValIleProLeuProGlnIleLeuGlyLeu 487
Db 4582 GTGAATCTTTTCAGAGAGGAAACTGAA----- 4608
QY 488 ThrAlaSerProGlyValGlyGlyAlaThrLysGlnAlaLysAla----- 502
Db 4609 -----ATATCAACCATTCAAAAGCAGTTAGAAAGCAATCAATGATAAATA 4653
QY 503 -----GluGluHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThrIleLys 519
Db 4654 CAGAACAAGATCCAAGAGATTATGAGAAAGAGGAAAGCAACTTAATATATAAACAAATAGT 4713

```

QY 520 ThrValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluProCysLysLysPhe 539
Db 4714 GAGGTTTCAGGAAACGTTGAATGAACCTGAAA---CAATTCAAGGAGCATCGCAA----- 4764
QY 540 AlaIleAlaAspAlaThrArgGluAspPropheLysGluLysLeuLeuGluIleMetThr 559
Db 4765 -----GCCAAGGATTTCAGCACTACAAAGTATAGAAAGTAAGATGCTCGAGTTGACCAAC 4818
QY 560 ArgIleGlnThrTyrCysGln-----MetSerProMetSerAspPheGlyThrGlnPro 577
Db 4819 AGACTTCAAGAAAGTCAAGAAAGAAATACAAATTATGATTAAAGGAAAAAGAGGAAATGAAA 4878
QY 578 TyrGluGlnTrpAlaIleGlnMetGluLysLysAlaAlaLysLysGlyAsnArg----- 595
Db 4879 AGAGTACAGGAGGCCCTTCAGATAGAGAGACCAACTGAAAGAAACACTAAAGAAATT 4938
QY 595 ----- 595
Db 4939 GTAGCTAAATGAAGAAATCTCAAGAAAAAGAAATATATCAGTTTCTTAAAGATGACAGCTGTC 4998
QY 596 -----LysGluArgValCys---AlaGluHisLeuArgLysTyrAsnGluAlaLeu 611
Db 4999 AATGAGACTCAGGAGAAATGTGTGAAATAGAACACTTGAAGGAGCAATTTGAGACCCAG 5058
QY 612 GlnIleAsn-----AspThrIleArgMetIleAsp---AlaTyrThr 624
Db 5059 AAGTTAAACCTGGAAAAACATAGAAAACGGAGAAATATAAGGTTGACTCAGATACATACATGAA 5118
QY 625 HisLeuGluThrPheTyrAsnGluGluLysAspLysLysPheAlaValIleGluAspAsp 644
Db 5119 AACCTTGAAGAAATGAGATCTGTAAACAAAGAAAGA-----GATGAC 5160
QY 645 SerAspGluGlyGlyAspAspGluTyrCysAspGlyAspGluAspGluAspLeuLys 664
Db 5161 CTTAGGAGTGTGGAGGAGACTCTCAAAGTAGAGAGAGACCAGCTCAAGGAAAAACCTTAGA 5220
QY 665 LysProLeuLysLeuAsp----- 670
Db 5221 GAAACTATAACTAGAGACCTAGAAAAACAAGAGGAGCTAAAAATTGTTCAATGTCATCTG 5280
QY 671 -----GluThrAspArgPheLeuMetThrLeuPhePheGluAsnAsnLysMet 686
Db 5281 AAGGAGCACCAAGAACTATTGTATAAACTAAGAGGGATTGTTTCAGAGAAAAACAATGAA 5340
QY 687 LeuLysArgLeuAlaGluAsnProGluTyrGluAsnGluLysLeu----- 701
Db 5341 ATATCAAATATGCAAAAGGACTTAGAACACTCAAATGATGCCTTAAAGACACAGGATCTG 5400
QY 702 -----ThrLysLeuArgAsnThrIleMetGluGlnTyrThrArgThrGlu 716
Db 5401 AAAATACAAGAGGAACCTAAGAATTGCTCACATGCATCTGAAAGAGCAGCAGGAACTATT 5460
QY 717 GluSerAlaArgGlyIleIlePheThrLysThrArgGlnSerAlaTyrAlaLeuSerGln 736
Db 5461 GACAAACTCAGAGGAATTGTTTCTGAGAAAGACAGATAAACTATCAAATATGCAAAAAAGAT 5520
QY 737 TrpIleThrGluAsn-----GluLysPheAlaGluValGlyValLysAlaHis 752
Db 5521 TTAGAAAAATTCAAATGCTAAATATACAAGAAAAAGATTCAAGAACTTAAGGCAAAATGAACAT 5580
QY 753 HisLeuIleGlyAlaGlyHisSer-----SerGluPheLys 764
Db 5581 CAACTTATTACGTTAAAAAAAAGATGTCAATGAGACACAGAAAAAAGTGTCTGAAATGGAG 5640
QY 765 ProMetThrGln-----AsnGluGlnLysGluValIleSerLysPheArgThrGlyLys 782
Db 5641 CAACTAAAGAAACAATAAAGACCACCAAGCTTAACCTCTGAGTAAATAGAAATAGAGAAT 5700
QY 783 IleAsnLeuLeuIleAlaThrThrValAlaGluGluGlyLeuAspIleLysGluCysAsn 802
Db 5701 TTAATTTTG-----GCTCAAGAACTTTCATGAAAAACCTTGAAGAAATGAAA----- 5745
QY 803 IleValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAlaArgGlyArg 822

Db 5746 ----- 5746
QY 823 AlaArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGlyValIle--- 841
Db 5770 CTAAGAAGAGTAGAGGAGACACTCAAACTGGAGAGAGACCAACTCAAGGAAAGCCTGCAA 5829
QY 841 ----- 841
Db 5830 GAAACCAAAGCTAGAGATCTGGAAATACAAACAGGAACTAAAACTGCTCGTATGCTATCA 5889
QY 842 ---GluHis---GluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHis 859
Db 5890 AAAGAACACAAAGAAACTGTTGATAAACTTAGAGAAAAAAATTTTCAGAAAAAGACAATTCAA 5949
QY 860 CysValGlnAsnMetLys-----ProGluGluTyrAlaHisLysIleLeu 874
Db 5950 ATTTTCAGACATTTCAAAAAGGATTTAGATAAAATCAAAAAGATGAATTACAGAAAAAGATCCAA 6009
QY 875 GluLeuGlnMetGlnSerIleMet-----Glu 883
Db 6010 GAACCTTCAGAAAAAGAACTTCAACTGCTTAGAGTGAAAGAGATGTCAATATGAGTCAT 6069
QY 884 LysLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAsnProSerLeuIle 903
Db 6070 AAAAAAATTAAATGAAATGGAACAGTTGAAGAAGCAATTTTGAGCCAAAC----- 6117
QY 904 ThrPheLeuCysLysAsnCysSerValLeuAlaCysSerGlyGluAspIleHisValIle 923
Db 6118 ---TATCTATGCAAG-----TGTGAGATGGATAAATTTCCAGTTTGACT 6156
QY 924 GluLysMetHisHisValAsnMetThrProGluPheLysGluLeuTyrIleValArgGlu 943
Db 6157 AAGAAACTTCATGAA-----AGCCTTGAAGAAATAAGAATTGTAGCTAAA 6201
QY 944 AsnLys 945
Db 6202 GAAAGA 6207

Search completed: January 28, 2005, 03:24:13
Job time : 3812 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 28, 2005, 00:36:18 ; Search time 1189 Seconds
(without alignments)
4953.352 Million cell updates/sec

Title: US-09-515-363C-2
Perfect score: 5311
Sequence: 1 MSGXGSTDENFRYLISCFRA.....LPITFPNLDYSECCLFSD 1025

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4300275 seqs, 2872944193 residues

Total number of hits satisfying chosen parameters: 8600550

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09515363/runat_26012005_152045_6240/app_query.fasta_1.1223
-DB=Published Applications NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09515363@cgn_1_1_723@runat_26012005_152045_6240
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
20: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
21: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

*

Result No.	Score	Query Match	Length	DB	ID	Description
1	5311	100.0	3365	14	US-10-055-475-1	Sequence 1, Appli
2	5311	100.0	3365	14	US-10-228-897-1	Sequence 1, Appli
3	5311	100.0	3380	17	US-10-755-889-405	Sequence 405, App
4	5307	99.9	3365	14	US-10-055-475-8	Sequence 8, Appli
5	5303	99.8	3928	14	US-10-198-846-13042	Sequence 13042, A
6	5299	99.8	3627	14	US-10-055-475-4	Sequence 4, Appli
7	5285	99.5	3372	16	US-10-275-822A-1	Sequence 1, Appli
8	5251	98.9	3668	18	US-10-723-860-5057	Sequence 5057, Ap
9	2874	54.1	1776	15	US-10-094-749-828	Sequence 828, App
10	2442	46.0	1443	16	US-10-275-822A-10	Sequence 10, Appl
11	2072	39.0	1392	16	US-10-641-643-48	Sequence 48, Appl
12	1995	37.6	1284	16	US-10-275-822A-7	Sequence 7, Appli
13	1352	25.5	2590	16	US-10-302-172-771	Sequence 771, App
14	1129	21.3	2250	16	US-10-108-260A-2072	Sequence 2072, Ap
15	1090.5	20.5	3065	15	US-10-172-118-1472	Sequence 1472, Ap
16	1090.5	20.5	3065	16	US-10-342-887-1472	Sequence 1472, Ap
17	1090.5	20.5	3065	17	US-10-755-889-233	Sequence 233, App
18	1090.5	20.5	3065	18	US-10-370-715B-7	Sequence 7, Appli
19	1090.5	20.5	3258	13	US-10-044-090-631	Sequence 631, App
c 20	783	14.7	499	9	US-09-864-761-29164	Sequence 29164, A
c 21	783	14.7	595	9	US-09-864-761-12563	Sequence 12563, A
22	782	14.7	609	15	US-10-106-698-811	Sequence 811, App
23	782	14.7	609	16	US-10-264-049-69	Sequence 69, Appl
c 24	777	14.6	453	15	US-10-029-386-24101	Sequence 24101, A
25	712	13.4	457	10	US-09-918-995-12953	Sequence 12953, A
26	666	12.5	1036	14	US-10-055-475-3	Sequence 3, Appli
27	666	12.5	1036	14	US-10-228-897-3	Sequence 3, Appli
28	666	12.5	6406	14	US-10-228-897-25	Sequence 25, Appl
29	563.5	10.6	3184	16	US-10-264-049-128	Sequence 128, App
30	553	10.4	403	14	US-10-198-846-5938	Sequence 5938, Ap
31	498	9.4	458	10	US-09-907-907A-32	Sequence 32, Appl
32	488	9.2	292	9	US-09-864-761-17369	Sequence 17369, A
33	472	8.9	301	15	US-10-417-827-12	Sequence 12, Appl
34	453	8.5	2259	15	US-10-369-493-45241	Sequence 45241, A
35	451.5	8.5	460	9	US-09-864-761-577	Sequence 577, App
36	441.5	8.3	1950	15	US-10-369-493-44088	Sequence 44088, A
37	441.5	8.3	1953	15	US-10-369-493-25025	Sequence 25025, A
38	431.5	8.1	5997	17	US-10-437-963-30018	Sequence 30018, A
39	417	7.9	2226	15	US-10-369-493-45047	Sequence 45047, A
40	408	7.7	5815	9	US-09-938-842A-887	Sequence 887, App
41	408	7.7	5815	11	US-09-938-842A-887	Sequence 887, App
42	395.5	7.4	2295	18	US-10-828-924-70	Sequence 70, Appl
c 43	382	7.2	554	9	US-09-864-761-15269	Sequence 15269, A
c 44	357	6.7	217	9	US-09-864-761-31791	Sequence 31791, A
45	357	6.7	10220	16	US-10-439-703-48	Sequence 48, Appl

ALIGNMENTS

RESULT 1
US-10-055-475-1
; Sequence 1, Application US/10055475
; Publication No. US20030022855A1
; GENERAL INFORMATION:
; APPLICANT: Fisher, Paul B.
; APPLICANT: Kang, Dong-Chul
; APPLICANT: Gopalkrishnan, Rahul V.
; TITLE OF INVENTION: USE OF MDA-5 AS AN ANTIVIRAL AND
; TITLE OF INVENTION: ANTIPROLIFERATIVE AGENT
; FILE REFERENCE: A34614-A-PCT-USA-A (070050.1921)
; CURRENT APPLICATION NUMBER: US/10/055,475
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: PCT/US01/06960
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 09/515,363
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3365

; TYPE: DNA			
; ORGANISM: homo sapiens			
US-10-055-475-1			
Alignment Scores:			
Pred. No.:	0	Length:	3365
Score:	5311.00	Matches:	1025
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	14	Gaps:	0
US-09-515-363C-2 (1-1025) x US-10-055-475-1 (1-3365)			
QY	1	MetSerAsnGlyTyrSerThrAspGluAsnPheArgTyrLeuIleSerCysPheArgAla	20
Db	169	ATGTCGAATGGGTATTCACAGACGAGAAATTTCCGCTATCTCATCTCGTCTCAGGGCC	228
QY	21	ArgValLysMetTyrIleGlnValGluProValLeuAspTyrLeuThrPheLeuProAla	40
Db	229	AGGGTGAATAATGTACATCCAGGTGGAGCCTGTGCTGGACTACCTGACCTTTCTGCCTGCA	288
QY	41	GluValLysGluGlnIleGlnArgThrValAlaThrSerGlyAsnMetGlnAlaValGlu	60
Db	289	GAGGTGAAGGAGCAGATTTCAGAGACAGTCGCCACCTCCGGGAACATGCAGCAGTTGAA	348
QY	61	LeuLeuLeuSerThrLeuGluLysGlyValTrpHisLeuGlyTrpThrArgGluPheVal	80
Db	349	CTGCTGCTGAGCACCTTGGAGAAGGAGTCTGGCACCTTGGTGGACTCGGGAATTCGTG	408
QY	81	GluAlaLeuArgArgThrGlySerProLeuAlaAlaArgTyrMetAsnProGluLeuThr	100
Db	409	GAGGCCCTCCGGAGAACCGGCAGCCTCTGGCCGCCGCTACATGAACCTGAGCTCACG	468
QY	101	AspLeuProSerProSerPheGluAsnAlaHisAspGluTyrLeuGlnLeuLeuAsnLeu	120
Db	469	GACTTGCCCTCTCCATCGTTTGAGAACGCTCATGATGAATATCTCCAACCTGCTGAACCTC	528
QY	121	LeuGlnProThrLeuValAspLysLeuLeuValArgAspValLeuAspLysCysMetGlu	140
Db	529	CTTCAGCCCACTCTGGTGGACAAAGCTTCTAGTTAGAGACGCTCTGGATAAGTGCATGGAG	588
QY	141	GluGluLeuLeuThrIleGluAspArgAsnArgIleAlaAlaAlaGluAsnAsnGlyAsn	160
Db	589	GAGGAACTGTTGACAATTTGAAGACAGAAACCGGATTGCTGCTGCAGAAACAAATGGAAAT	648
QY	161	GluSerGlyValArgGluLeuLeuLysArgIleValGlnLysGluAsnTrpPheSerAla	180
Db	649	GAATCAGGTGTAAGAGAGCTACTATAAAGGATTGTGCAGAAAGAAACTGGTTCTCTGCA	708
QY	181	PheLeuAsnValLeuArgGlnThrGlyAsnAsnGluLeuValGlnGluLeuThrGlySer	200
Db	709	TTTCTGAATGTTCTTCGTCAAACAGGAAACAATGAACCTTGCTCAAGAGTTAACAGGCTCT	768
QY	201	AspCysSerGluSerAsnAlaGluIleGluAsnLeuSerGlnValAspGlyProGlnVal	220
Db	769	GATTGCTCAGAAAGCAATGCAGAGATTGAGAATTTATCACAAAGTTGATGGTCTCCTCAAGTG	828
QY	221	GluGluGlnLeuLeuSerThrThrValGlnProAsnLeuGluLysGluValTrpGlyMet	240
Db	829	GAAGAGCAACTTCTTCAACCACAGTTCGCCAAATCTGGAGAAGGAGGTCCTGGGCGCATG	888
QY	241	GluAsnAsnSerSerGluSerPheAlaAspSerSerValValSerGluSerAspThr	260
Db	889	GAGAATAACTCATCAGAAATCATCTTTTGCAGATCTTCTGTAGTTTCAGAAATCAGACACA	948
QY	261	SerLeuAlaGluGlySerValSerCysLeuAspGluSerLeuGlyHisAsnSerAsnMet	280
Db	949	AGTTTGGCAGAAAGGAAGTGTGAGTGAAGTCTTGGACATAACAGCAACATG	1008
QY	281	GlySerAspSerGlyThrMetGlySerAspSerAspGluGluAsnValAlaAlaArgAla	300
Db	1009	GGCAGTGATTTCAGGCACCATGGGAAGTGATTTCAGATGAAGAGAAATGTGGCAGCAAGAGCA	1068
QY	301	SerProGluProGluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaLeu	320
Db	1069	TCCCCGGAGCCAGAACTCCAGCTCAGGCCCTTACCAATGGAAGTTGCCAGCAGCCTTG	1128
QY	321	GluGlyLysAsnIleIleCysLeuProThrGlySerGlyLysThrArgValAlaVal	340
Db	1129	GAAGGGAAGAATATCATCATCTGCCTCCCTACAGGGAGTGGAATAAACAGAGTGGCTGTT	1188
QY	341	TyrIleAlaLysAspHisLeuAspLysLysLysLysAlaSerGluProGlyLysValIle	360
Db	1189	TACATTGCCAAGGATCACCTTAGACAAGAGAAAAAAGCATCTGAGCCTGGAAGATTATA	1248
QY	361	ValLeuValAsnLysValLeuValGluGlnLeuPheArgLysGluPheGlnProphe	380
Db	1249	GTTCTTGTCAATAAGGTACTGCTAGTTGAACAGCTCTTCGCAAGGAGTTCCACCATTT	1308
QY	381	LeuLysLysTrpTyrArgValIleGlyLeuSerGlyAspThrGlnLeuLysIleSerPhe	400
Db	1309	TTGAAGAAATGGTATCGTGTATTGATTAAGTGGTGATACCCAACTGAAATATCATTT	1368
QY	401	ProGluValValLysSerCysAspIleIleIleSerThrAlaGlnIleLeuGluAsnSer	420
Db	1369	CCAGAAGTTGTCAAGTCTCTGTATATTTATATCAGTACAGTCAAACTCCTTGAAAACTCC	1428
QY	421	LeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPheSerLeuIle	440
Db	1429	CTCTTAAACTTGGAAAAATGGAGAAGATGCTGGTGTTCATTGTTCAGACTTTTCCCTCAT	1488
QY	441	IleIleAspGluCysHisHisThrAsnLysGluAlaValTyrAsnAsnIleMetArgHis	460
Db	1489	ATCATTTGATGATGTCATCACACCAACAAAGAGCAGTGTATAATAACATCATGAGGCAT	1548
QY	461	TyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysLysGluAsnLysProValIlePro	480
Db	1549	TATTTGATGCAGAGATTGAAAAACAATAGACTCAAGAAAGAAAAACAACCAAGTATTCCC	1608
QY	481	LeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyGlyAlaThrLysGlnAla	500
Db	1609	CTTCTCAGATACTGGGACTAACAGCTTACCTGGTGTGGAGGGGCCACGAAGCAAGCC	1668
QY	501	LysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThrIleLysThr	520
Db	1669	AAAGCTGAAGAACACATTTTAAACCTATGTGCCAATCTTGTGATGCTATTATAAACT	1728
QY	521	ValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluProCysLysLysPheAla	540
Db	1729	GTTAAAGAAAAACCTTGATCAACTGAAAAAACCAATAACAGGAGCCATGCAAGAGTTTGGC	1788
QY	541	IleAlaAspAlaThrArgGluAspProPheLysGluLysLeuLeuIleMetThrArg	560
Db	1789	ATTGCAGATGCAACCAGAGAAGATCCATTAAAGAGAGAACTTTAGAAAAATAATGACAAGG	1848
QY	561	IleGlnThrTyrCysGlnMetSerProMetSerAspPheGlyThrGlnProTyrGluGln	580
Db	1849	ATTCAAACTTATTGTCAAATGAGTCCAATGTCCAGATTTTGGAACTCAACCCCTATGAACAA	1908
QY	581	TrpAlaIleGlnMetGluLysLysAlaAlaLysLysGlyAsnArgLysGluArgValCys	600
Db	1909	TGGCCATTCAAATGGAATAAAAAAGCTGCAAAAAAGGAAATCGCAAAAGAACGTTGTTGT	1968
QY	601	AlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMetIle	620
Db	1969	GCAGAACATTTGAGGAAGTACAATGAGGCCCTTACAAATTAATGACACAATTCGAATGATA	2028
QY	621	AspAlaTyrThrHisLeuGluThrPheTyrAsnGluGluLysAspLysLysPheAlaVal	640
Db	2029	GATGCGTACTACTCATCTTGAAACTTTCTATATGAAGAGAAAGATAAGAGTTTGCAGTC	2088
QY	641	IleGluAspAspSerAspGluGlyGlyAspAspGluTyrCysAspGlyAspGluAspGlu	660
Db	2089	ATAGAAGATGATAGTATGAGGGTGGTGTATGATGATGATGATGATGATGATGATGATGATG	2148

```
QY 661 AspAspLeuLysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeuPhe 680
Db 2149 GATGATTTAAAGAAACCTTTGAAACTGGATGAAACAGATAGATTTCTCATGACTTTATTT 2208

QY 681 PheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGluLys 700
Db 2209 TTTGAAAAACAATAAAATGTTGAAAAGGCTGGCTGAAAACCCAGAAATATGAAAATGAAAAAG 2268

QY 701 LeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArgThrGluGluSerAlaArg 720
Db 2269 CTGACCAATTAAGAAATACCATATGGAGCAATATACTAGGACTGAGGAATCAGCACGA 2328

QY 721 GlyIleIlePheThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTrpIleThrGlu 740
Db 2329 GGAATAATCTTTACAAAAACACGACAGAGTGCATATGCGCTTTCCAGTGGATTACTGAA 2388

QY 741 AsnGluLysPheAlaGluValGlyValLysAlaHisHisLeuIleGlyAlaGlyHisSer 760
Db 2389 AATGAAAAATTTGCTGAAGTAGGAGTCAAAGCCCCACCATCTGTATTGGAGCTGGACACAGC 2448

QY 761 SerGluPheLysProMetThrGlnAsnGluGlnLysGluValIleSerLysPheArgThr 780
Db 2449 AGTGAGTTCAAACCCCATGACACAGAATGAACAAAAAGAAAGTCAATTAGTAAATTTCCGCACT 2508

QY 781 GlyLysIleAsnLeuLeuIleAlaThrThrValAlaGluGluGlyLeuAspIleLysGlu 800
Db 2509 GGAAAAATCAATCTGCTTATCGCTACCACAGTGGCAGAGAAGGTCTGGATATTAAAGAA 2568

QY 801 CysAsnIleValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAlaArg 820
Db 2569 TGTAACATTGTTATCCGTTATGGTCTCGTCACCAATGAATAGCCATGGTCCAGGCCCGT 2628

QY 821 GlyArgAlaArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGlyVal 840
Db 2629 GGTGAGCCAGAGCTGATGAGAGCACCTACGTCTGTTGCTCACAGTGGTTCAGGAGTT 2688

QY 841 IleGluHisGluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHisCys 860
Db 2689 ATCGAACATGAGACAGTTAATGATTTCCGAGAGAAGATGATGTATAAGCTATACATTGT 2748

QY 861 ValGlnAsnMetLysProGluGluTyrAlaHisLysIleLeuGluLeuGlnMetGlnSer 880
Db 2749 GTTCAAATATGAAACCCAGAGGAGTATGCTCATAAGATTTTGGAAATTACAGATGCCAAAGT 2808

QY 881 IleMetGluLysLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAsnPro 900
Db 2809 ATAATGAAAAAGAAATGAAACCAAGAGAAATATTGCCAAGCATTTACAAGATAACCCA 2868

QY 901 SerLeuIleThrPheLeuCysLysAsnCysSerValLeuAlaCysSerGlyGluAspIle 920
Db 2869 TCACATAATAACTTTCCTTTTGCAAAAACTGCAGTGTGCTAGCCTGTTCTGGGGAAGATATC 2928

QY 921 HisValIleGluLysMetHisHisValAsnMetThrProGluPheLysGluLeuTyrIle 940
Db 2929 CATGTAATTGAGAAAATGCATCACGTCAATATGACCCAGAAATTCAGGAACTTTACATT 2988

QY 941 ValArgGluAsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluIle 960
Db 2989 GTAAGAGAAAACAAAGCACCTGCAAAAAGAAGTGTCCGCACTATCAAATAAATGGTGAAATC 3048

QY 961 IleCysLysCysGlyGlnAlaTrpGlyThrMetMetValHisLysGlyLeuAspLeuPro 980
Db 3049 ATCTGCAATGTGGCCAGGCTTGGGGAACAATGATGGTGCACAAAGGCTTAGATTGTCCT 3108

QY 981 CysLeuLysIleArgAsnPheValValPheLysAsnAsnSerThrLysLysGlnTyr 1000
Db 3109 TGTCTCAAATAAGGAATTTTGTAGTGGTTTTCAAAAAATAATTCAACAAAGAAACAATAC 3168

QY 1001 LysLysTrpValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCysLysLeu 1020
Db 3169 AAAAAAGTGGGTAGAAATTACCTATCACATTTCCCAATCTTGACTATTTCAGAAATGCTGTTA 3228

QY 1021 PheSerAspGluAsp 1025
```

```
Db 3229 TTTTAGTGATGAGGAT 3243

RESULT 2
US-10-228-897-1
; Sequence 1, Application US/10228897
; Publication No. US20030092043A1
; GENERAL INFORMATION:
; APPLICANT: Fisher, Paul B.
; APPLICANT: Kang, Dong-Chul
; APPLICANT: GopalKrishnan, Rahul V.
; TITLE OF INVENTION: MELANOMA DIFFERENTIATION ASSOCIATED
; TITLE OF INVENTION: GENE-5 AND PROMOTER AND USES THEREOF
; FILE REFERENCE: A34614-A-PCT-USA (070050.2121)
; CURRENT APPLICATION NUMBER: US/10/228,897
; CURRENT FILING DATE: 2002-08-26
; PRIOR APPLICATION NUMBER: PCT/US01/06960
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 09/515,363
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3365
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-228-897-1

Alignment Scores:
Pred. No.: 0 Length: 3365
Score: 5311.00 Matches: 1025
Percent Similarity: 100.00% Conservative: 0
Best local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-09-515-363C-2 (1-1025) x US-10-228-897-1 (1-3365)

QY 1 MetSerAsnGlyTyrSerThrAspGluAsnPheArgTyrLeuIleSerCysPheArgAla 20
Db 169 ATGTCGAATGGGTATTCACAGACAGAGAAATTTCCGCTATCTCATCTCGTGTTCAGGGCC 228

QY 21 ArgValLysMetTyrIleGlnValGluProValLeuAspTyrLeuThrPheLeuProAla 40
Db 229 AGGTTGAAAATGTACATCCAGGTGGAGCCTGTGCTGGACTACCTGACCTTTCTGCTGCA 288

QY 41 GluValLysGluGlnIleGlnArgThrValAlaThrSerGlyAsnMetGlnAlaValGlu 60
Db 289 GAGGTGAAGGAGCAGATTTCAGAGGACAGTCGCCACCTCCGGGAACATGCAGGCACTTGAA 348

QY 61 LeuLeuLeuSerThrLeuGluLysGlyValTrpHisLeuGlyTrpThrArgGluPheVal 80
Db 349 CTGCTGCTGAGCACCTTGGAGAAGGGAGTCTGGCACCTTGGTTGGACTCGGGAATTCGTG 408

QY 81 GluAlaLeuArgThrGlySerProLeuAlaIleArgTyrMetAsnProGluLeuThr 100
Db 409 GAGGCCCTCCGAGAACCCGCGAGCCCTCTGGCCGCCGCTACATGAACCCCTGAGCTCACG 468

QY 101 AspLeuProSerPheGluAsnAlaHisAspGluTyrLeuGlnLeuLeuAsnLeu 120
Db 469 GACTTGCCCTCTCCATCGTTTGAGAACGCTCATGATGAATATCTCCAACCTGCTGAACCTC 528

QY 121 LeuGlnProThrLeuValAspLysLeuLeuValArgAspValLeuAspLysCysMetGlu 140
Db 529 CTTCAGCCCACTCTGGTGGACAAGCTTCTAGTTAGACAGCTCTTTGGATAAGTCATGGAG 588

QY 141 GluGluLeuLeuThrIleGluAspArgAsnArgIleAlaAlaGluAsnAsnGlyAsn 160
Db 589 GAGGAACCTGTTGACAAATTGAAGACAGAAACCCGGATTGCTGCTGCAGAAAAACAATGGAAT 648

QY 161 GluSerGlyValArgGluLeuLeuLysArgIleValGlnLysGluAsnTrpPheSerAla 180
Db 649 GAATCAGGTGTAAGAGAGCTACTAAAAAAGGATTGTGCAGAAAAAGAACTGGTTCTCTGCA 708
```

Qy	181	pheLeuAsnValLeuArgGlnThrGlyAsnAsnGluLeuValGlnGluLeuThrGlySer	200	Qy	541	IleAlaAspAlaThrArgGluAspProPheLysGluLysLeuLeuGluIleMetThrArg	560
Db	709	TTTCTGAATGTTCTTCGTCAAACAGGAAACAATGAACCTGTCCAAGAGTTAAACAGGCTCT	768	Db	1789	ATTGCAGATGCAACCAAGAGATCCATTAAAGAGAAACTTCTAGAAATAATGACAAGG	1848
Qy	201	AspCysSerGluSerAsnAlaGluIleGluAsnLeuSerGlnValAspGlyProGlnVal	220	Qy	561	IleGlnThrTyrCysGlnMetSerProMetSerAspPheGlyThrGlnProTyrGluGln	580
Db	769	GATTGCTCAGAAAGCAATGCAGAGATTGAGAAATTTATCAACAAGTTGATGGTCCCTCAAGTG	828	Db	1849	ATTCAAACTTATTGTCAAAATGAGTCCAATGTCCAGATTTTGGAACCTCAACCCCTATGAACAA	1908
Qy	221	GluGluGlnLeuLeuSerThrThrValGlnProAsnLeuGluLysGluValTrpGlyMet	240	Qy	581	TrpAlaIleGlnMetGluLysLysAlaAlaLysLysGlyAsnArgLysGluArgValCys	600
Db	829	GAAGAGCAACTCTTTCAACCACAGTTCAGCCAAATCTGGAGAAGGAGTCTGGGGCATG	888	Db	1909	TGGCCATTCAAATGGAAAAAAGCTGCAAAAAAGGAAATCGCAAAAGAACGTGTTTGT	1968
Qy	241	GluAsnAsnSerSerGluSerSerPheAlaAspSerSerValValSerGluSerAspThr	260	Qy	601	AlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMetIle	620
Db	889	GAGAATAACTCATCAGAAATCATCTTTTGCAGATTCCTTCGTAGTTTCAGAATCAGACACA	948	Db	1969	GCAGAACATTTGAGGAAGTACAATGAGGCCCTACAAATTAATGACACAATTCGAATGATA	2028
Qy	261	SerLeuAlaGluGlySerValSerCysLeuAspGluSerLeuGlyHisAsnSerAsnMet	280	Qy	621	AspAlaTyrThrHisLeuGluThrPheTyrAsnGluLysAspLysLysPheAlaVal	640
Db	949	AGTTTGGCAGAGGAAGTGTCCAGCTGCTTAGATGAAAGTCTTGGACATAACAGCAACATG	1008	Db	2029	GATGCGTATACTCATCTTGAAACTTTCTATAATGAAGAGAAAGATAAGAAAGTTTGCAGTC	2088
Qy	281	GlySerAspSerGlyThrMetGlySerAspSerAspGluGluAsnValAlaAlaArgAla	300	Qy	641	IleGluAspAspSerAspGluGlyGlyAspAspGluTyrCysAspGlyAspGluAspGlu	660
Db	1009	GGCAGTGATTCAGGCACCATGGGAAGTGATTGAGATGAAGAAATGTGGCAGCAAGAGCA	1068	Db	2089	ATAGAAGATGATAGTGATGAGGGTGGTGATGATGAGTATTGTGTGGTGGATGAAGATGAG	2148
Qy	301	SerProGluProGluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaLeu	320	Qy	661	AspAspLeuLysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeuPhe	680
Db	1069	TCCCCGGAGCCAGAACTCCAGCTCAGGCCTTACCAAATGGAAGTTGCCAGCCAGCCTTG	1128	Db	2149	GATGATTTAAAGAAACCTTTGAAACTGGATGAAACAGATAGATTTCCTCATGACTTTATTT	2208
Qy	321	GluGlyLysAsnIleIleCysLeuProThrGlySerGlyLysThrArgValAlaVal	340	Qy	681	PheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGluLys	700
Db	1129	GAAGGGAAGATATCATCATCTGCCTCCCTACAGGGAGTGGAAAAACCAGAGTGGCTGTT	1188	Db	2209	TTTGAAAAACAATAAAATGTTGAAAAGGCTGGCTGAAAAACCAGAAATATGAAAATGAAAAG	2268
Qy	341	TyrIleAlaLysAspHisLeuAspLysLysLysAlaSerGluProGlyLysValIle	360	Qy	701	LeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArgThrGluGluSerAlaArg	720
Db	1189	TACATTGCCAAGGATCACTTAGACAAGAAAAAAGCATCTGAGCCTGGAAAAAGTTATA	1248	Db	2269	CTGACCAAAATTAAGAAATACCATTAATGGAGCAATATACTAGGACTGAGGAATCAGCACGA	2328
Qy	361	ValLeuValAsnLysValLeuLeuValGluGlnLeuPheArgLysGluPheGlnProPhe	380	Qy	721	GlyIleIlePheThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTrpIleThrGlu	740
Db	1249	GTTCTTGTCAATAAGGTACTGTCTAGTTGAACAGCTCTTCCGCAAGGAGTTCCAACCATTT	1308	Db	2329	GGAATAATCTTTACAAAAACACGACAGAGTGCATATGCGCTTTCCAGTGGATTACTGAA	2388
Qy	381	LeuLysLysTrpTyrArgValIleGlyLeuSerGlyAspThrGlnLeuLysIleSerPhe	400	Qy	741	AsnGluLysPheAlaGluValGlyValLysAlaHisHisLeuIleGlyAlaGlyHisSer	760
Db	1309	TTGAAGAAATGGTATCGTGTATTGGATTAAAGTGGTGATACCCAACTGAAAAATATCATTT	1368	Db	2389	AATGAAAAATTTGCTGAAGTAGAGTCAAAGCCCAACCATCTGATTGGAGCTGGACACAGC	2448
Qy	401	ProGluValValLysSerCysAspIleIleIleSerThrAlaGlnIleLeuGluAsnSer	420	Qy	761	SerGluPheLysProMetThrGlnAsnGluGlnLysGluValIleSerLysPheArgThr	780
Db	1369	CCAGAAGTTGTCAAGTCTGTGATATTATTATCAGTACAGTCAAAATCCTTGAAAACTCC	1428	Db	2449	AGTGAGTTCAAACCCCATGACACAGAATGAACAAAGAGAGTCAATAGTAAATTCGCACT	2508
Qy	421	LeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPheSerLeuIle	440	Qy	781	GlyLysIleAsnLeuLeuIleAlaThrThrValAlaGluGluGlyLeuAspIleLysGlu	800
Db	1429	CTCTTAAACTTGGAAAAATGGAGAAGATGCTGGTGTTCAAATTGTCCAGACTTTTCCCTCAT	1488	Db	2509	GGAAAAATCAATCTGCTTATCGTACCACAGTGGCAGAGAAGAGTCTGGATATTAAGAA	2568
Qy	441	IleIleAspGluCysHisHisThrAsnLysGluAlaValTyrAsnAsnIleMetArgHis	460	Qy	801	CysAsnIleValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAlaArg	820
Db	1489	ATCATTTGATGAATGTATCATCACCAACAAAGAACAGCAGTGTATAATAACATCATGAGGCAT	1548	Db	2569	TGTAACATTTGTTATCCGTTATGGTCTCGTCACCAATGAATAGCCATGGTCCAGGCCGT	2628
Qy	461	TyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysLysGluAsnLysProValIlePro	480	Qy	821	GlyArgAlaArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGlyVal	840
Db	1549	TATTTGATGCAGAAAGTTGAAAAACAATAGACTCAAGAAAAAGAAAAACAAACCCAGTATCCC	1608	Db	2629	GGTCGAGCCAGAGCTGATGAGAGCACCTACGTCCTGGTTGCTCACAGTGGTTCAGGAGTT	2688
Qy	481	LeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyGlyAlaThrLysGlnAla	500	Qy	841	IleGluHisGluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHisCys	860
Db	1609	CTTCTTCAGATACTGGGACTAACAGCTTCACCTGGTGTGGAGGGGCCACGAAGCAAGCC	1668	Db	2689	ATCGAACATGAGACAGTTAATGATTTCCGAGAGAGATGATGTATAAAGCTATACATTGT	2748
Qy	501	LysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThrIleLysThr	520	Qy	861	ValGlnAsnMetLysProGluGluTyrAlaHisLysIleLeuGluLeuGlnMetGlnSer	880


```

Db      2869 TCACATAAATCTTCCTTTGCAAAAACGTCAGTGTGCTAGCCTGTTCTGGGAAGATATC 2928
QY      921 HisValIleGluLysMetHisHisValAsnMetThrProGluPheLysGluLeuTyrIle 940
Db      2929 CATGTAATTGAGAAAATGCATCACGTCAAATATGACCCAGAAATTCAGGAACTTTACATT 2988
QY      941 ValArgGluAsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluIle 960
Db      2989 GTAAGAGAAAAACAAAGCACTGCAAAAGAAAGTGTGCCGACTATCAAAATAAATGGTGAATC 3048
QY      961 IleCysLysCysGlyGlnAlaTrpGlyThrMetMetValHisLysGlyLeuAspLeuPro 980
Db      3049 ATCTGCAAAATGTGCCAGGCTTGGGGAACAATGATGGTGCAAAAGGCTTAGATTGCTT 3108
QY      981 CysLeuLysIleArgAsnPheValValPheLysAsnAsnSerThrLysLysGlnTyr 1000
Db      3109 TGTCTCAAAATAAGGAATTTTGTAGTGGTTTCAAAAATAATTCAACAAAGAAACAATAC 3168
QY      1001 LysLysTrpValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCysLysLeu 1020
Db      3169 AAAAAGTGGTAGAATTACCTATCACATTTCCCAATCTTGACTATTTCAGAAATGCTGTTTA 3228
QY      1021 PheSerAspGluAsp 1025
Db      3229 TTTAGTGATGAGGAT 3243
```

```

RESULT 3
US-10-755-889-405
; Sequence 405, Application US/10755889
; Publication No: US20040171823A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-kB
; TITLE OF INVENTION: PATHWAY
; FILE REFERENCE: D0284 NP
; CURRENT APPLICATION NUMBER: US/10/755,889
; CURRENT FILING DATE: 2004-01-13
; PRIOR APPLICATION NUMBER: U.S. 60/440,068
; PRIOR FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: U.S. 60/469,757
; PRIOR FILING DATE: 2003-05-12
; NUMBER OF SEQ ID NOS: 823
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 405
; LENGTH: 3380
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-755-889-405
```

```

Alignment Scores:
Pred. No.:      0      Length:      3380
Score:          5311.00  Matches:      1025
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match:      100.00%  Indels: 0
DB:               17      Gaps: 0
```

```

US-09-515-363C-2 (1-1025) x US-10-755-889-405 (1-3380)

QY      1 MetSerAsnGlyTyrSerThrAspGluAsnPheArgTyrLeuIleSerCysPheArgAla 20
Db      169 ATGTGCAATGGGTATTCACAGACGAGAAATTTCCGCTATCTCATCTCGTCTTCAGGGCC 228
QY      21 ArgValLysMetTyrIleGlnValGluProValLeuAspTyrLeuThrPheLeuProAla 40
Db      229 AGGGTGAAAATGTACATCCAGGTGGAGCCTGTGCTGGACTACTGACCTTTCTGCTGCA 288
QY      41 GluValLysGluGlnIleGlnArgThrValAlaThrSerGlyAsnMetGlnAlaValGlu 60
Db      289 GAGGTGAAGGAGCAGATTTCAGAGGACATCGCCACCTCCGGGAACATGCAGGCAGTTGAA 348
QY      61 LeuLeuLeuSerThrLeuGluLysGlyValTrpHisLeuGlyTrpThrArgGluPheVal 80
```

```

Db      349 CTGCTGCTGAGCACCTTGGAGAAGGAGTCTGGCACCTTGGTTGGACTCGGGAATTCGTG 408
QY      81 GluAlaLeuArgArgThrGlySerProLeuAlaAlaArgTyrMetAsnProGluLeuThr 100
Db      409 GAGGCCCTCCGGAGAACCGGGACGCCCTCTGGCCGCCCGCTACATGAACCCCTGAGCTCAGC 468
QY      101 AspLeuProSerProSerPheGluAsnAlaHisAspGluTyrLeuGlnLeuLeuAsnLeu 120
Db      469 GACTTGCCCTCTCCATCGTTTTGAGAACGCTCATGATGAATATCTCCAACCTGCTGAACCTC 528
QY      121 LeuGlnProThrLeuValAspLysLeuLeuValArgAspValLeuAspLysCysMetGlu 140
Db      529 CTTCAGCCCACTCTGGTGGACAAGCTTCTAGTTAGAGAGCTCTTGGATAAGTGCATGGAG 588
QY      141 GluGluLeuLeuThrIleGluAspArgAsnArgIleAlaAlaGluAsnAsnGlyAsn 160
Db      589 GAGGAACCTGTGACAAATTGAAGACAGAAACCGGATTGCTGCTGCAGAAAACAATGGAAAT 648
QY      161 GluSerGlyValArgGluLeuLeuLysArgIleValGlnLysGluAsnTrpPheSerAla 180
Db      649 GAATCAGGTGTAAAGAGAGCTACTAAAAGGATTGTGCAGAAAGAAACTGGTTCTCTGCA 708
QY      181 PheLeuAsnValLeuArgGlnThrGlyAsnAsnGluLeuValGlnGluLeuThrGlySer 200
Db      709 TTTCTGAATGTTCTTCGTCAAACAGGAAAACAATGAACTTGTCCAAGAGTTAACAGGCTCT 768
QY      201 AspCysSerGluSerAsnAlaGluIleGluAsnLeuSerGlnValAspGlyProGlnVal 220
Db      769 GATTGCTCAGAAAAGCAATGCAGAGATTGAGAAATTTATCACAAGTTGATGTCCTCAAGTG 828
QY      221 GluGluGlnLeuLeuSerThrThrValGlnProAsnLeuGluLysGluValTrpGlyMet 240
Db      829 GAAGAGCAACTTCTTTCAACCACAGTTCAGCCAAATCTGGAGAAGGAGGTCTGGGGCATG 888
QY      241 GluAsnAsnSerSerGluSerSerPheAlaAspSerSerValValSerGluSerAspThr 260
Db      889 GAGAATAACTCATCAGAATCATCTTTTGCAGATTCTTCTGTAGTTTCAGAAATCAGACACA 948
QY      261 SerLeuAlaGluGlySerValSerCysLeuAspGluSerLeuGlyHisAsnSerAsnMet 280
Db      949 AGTTTGGCAGAAGGAAGTGTACGCTGCTTAGATGAAAGTCTTGGACATAACAGCAACATG 1008
QY      281 GlySerAspSerGlyThrMetGlySerAspSerAspGluGluAsnValAlaAlaArgAla 300
Db      1009 GGCAGTGATTTCAGGCACCATGGGAAGTGATTTCAGATGAAGAGAATGTGGCAGCAAGAGCA 1068
QY      301 SerProGluProGluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaLeu 320
Db      1069 TCCCCGGAGCCAGAACTCCAGCTCAGGCCTTACCAAATGGAAGTTGCCAGCCAGCCTTG 1128
QY      321 GluGlyLysAsnIleIleCysLeuProThrGlySerGlyLysThrArgValAlaVal 340
Db      1129 GAAGGGAAGAATATCATCATCTGCCTCCCTACAGGGAGTGGAAAAACCCAGAGTGGCTGTT 1188
QY      341 TyrIleAlaLysAspHisLeuAspLysLysLysLysAlaSerGluProGlyLysValIle 360
Db      1189 TACATTGCCAAGGATCACTTAGACAAGAGAAAAAAGCATCTGAGCCTGGAAAAAGTTATA 1248
QY      361 ValLeuValAsnLysValLeuLeuValGluGlnLeuPheArgLysGluPheGlnProPhe 380
Db      1249 GTTCTTGTCAATAAGGTACTGCTAGTTGAACAGCTCTTCGCAAGGAGTTTCCAACCATTT 1308
QY      381 LeuLysLysTrpTyrArgValIleGlyLeuSerGlyAspThrGlnLeuLysIleSerPhe 400
Db      1309 TTGAAGAAATGGTATCGTGTATTGGATTAAAGTGGTGATACCCCAACTGAAAAATATCATTT 1368
QY      401 ProGluValValLysSerCysAspIleIleIleSerThrAlaGlnIleLeuGluAsnSer 420
Db      1369 CCAGAAGTTGTCAAGTCCTCTGTATATTATTATCAGTACAGCTCAAATCCTTTGAAAACTCC 1428
QY      421 LeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPheSerLeuIle 440
```


Db 1429 CTCTTAAACTTGGAAAAATGGAGAAGATGCTGGTGTTCATATGTCAGACTTTTCCCTCATT 1488

QY 441 IleIleAspGluCysHisHisThrAsnLysGluAlaValTyrAsnAsnIleMetArgHis 460

Db 1489 ATCATTGATGAATGTCATCACACCAACAAAGACAGTGTATATAATACATCATGAGGCAT 1548

QY 461 TyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysLysGluAsnLysProValIlePro 480

Db 1549 TATTTGATGCAGAAAGTTGAAAAACAATAGACTCAAGAAAGAAAAACAAACCAAGTATTCCC 1608

QY 481 LeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyGlyAlaThrLysGlnAla 500

Db 1609 CTTCCCTCAGATACTGGGACTAACAGCTTCACCTGGTGTGGAGGGGCCACGAAGCAAGCC 1668

QY 501 LysAlaGluCysHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThrIleLysThr 520

Db 1669 AAAGCTGAAGAACACATTTTAAACTATGTGCCAATCTTGATGCATTTACTATTAAACT 1728

QY 521 ValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluProCysLysLysPheAla 540

Db 1729 GTTAAAGAAACCTTGATCAACTGAAAAACCAATATACAGAGCCATGCAAGAGTTTGCC 1788

QY 541 IleAlaAspAlaThrArgGluAspProPheLysGluLysLeuGluIleMetThrArg 560

Db 1789 ATTGCAGATGCAACACAGAGAGATCCATTTAAAGAGAAACTTCTAGAAATAATGACAAGG 1848

QY 561 IleGlnThrTyrCysGlnMetSerProMetSerAspPheGlyThrGlnProTyrGluGln 580

Db 1849 ATTCAAACTTATGTCAATGAGTCCAATGTCCAGATTTTGGAACTCAACCCATGAAACA 1908

QY 581 TrpAlaIleGlnMetGluLysLysAlaAlaLysLysGlyAsnArgLysGluArgValCys 600

Db 1909 TGGGCCATTCAAATGGAAAAAAGCTGCAAAAAAGGAAATCGCAAGAACGTGTTGT 1968

QY 601 AlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMetIle 620

Db 1969 GCAGAACATTTGAGGAAGTACAAATGAGCCCTACAAATTAATGACACAAATTCGAATGATA 2028

QY 621 AspAlaTyrThrHisLeuGluThrPheTyrAsnGluGluLysAspLysLysPheAlaVal 640

Db 2029 GATCGGTATACTCATCTTGAAACTTTCTATAATGAAGAGAAAGATAAGAAAGTTTGCAGTC 2088

QY 641 IleGluAspAspSerAspGluGlyGlyAspAspGluTyrCysAspGlyAspGluAspGlu 660

Db 2089 ATAGAAGATGATAGTATGATGAGGGTGGTATGATGAGTATTTGTGATGATGAGATGAG 2148

QY 661 AspAspLeuLysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeuPhe 680

Db 2149 GATGATTTTAAAGAAACCTTTGAAACTGGATGAAACAGATAGATTTCTCATGACTTTATT 2208

QY 681 PheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGluLys 700

Db 2209 TTTGAAAAACAATAAATGTTGAAAAAGCTGGCTGAAAAACCCAGAAATATGAAAAAG 2268

QY 701 LeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArgThrGluGluSerAlaArg 720

Db 2269 CTGACCAATTAAGAAATACCATATGGAGCAATATACTAGGACTGAGGAATCAGCACGA 2328

QY 721 GlyIleIlePheThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTrpIleThrGlu 740

Db 2329 GGAATAATCTTTACAAAAACACGACAGAGTGCATATGCGCTTCCCAGTGGATTACTGAA 2388

QY 741 AsnGluLysPheAlaGluValGlyValLysAlaHisHisLeuIleGlyAlaGlyHisSer 760

Db 2389 AATGAAAAATTTGCTGAAGTAGGAGTCAAAGCCCATCTGATTGGAGCTGGACACAGC 2448

QY 761 SerGluPheLysProMetThrGlnAsnGluGlnLysGluValIleSerLysPheArgThr 780

Db 2449 AGTGAGTTCAAACCCATGACACAGAAATGAACAAAAAGAGTCATTAGTAAATTCGCACT 2508

QY 781 GlyLysIleAsnLeuLeuIleAlaThrThrValAlaGluGluGlyLeuAspIleLysGlu 800

Db 2509 GGAAAAATCAATCTGTCTTATCGCTACCACAGTGGCAGAAAGGCTTGGATATTAAGAA 2568

QY 801 CysAsnIleValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAlaArg 820

Db 2569 TGTAACATTGTTATCCGTTATGGTCTCGTCACCAATGAAATAGCCATGGTCCAGGCCGT 2628

QY 821 GlyArgAlaArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGlyVal 840

Db 2629 GGTGAGCCAGAGCTGATGAGAGCACCTACGTCCTGGTTGCTCACAGTGGTTCCAGAGTT 2688

QY 841 IleGluHisGluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHisCys 860

Db 2689 ATCGAACATGAGACAGTTAATGATTTCCGAGAGAGATGATGTATATAAGCTATACATTGT 2748

QY 861 ValGlnAsnMetLysProGluGluTyrAlaHisLysIleLeuGluLeuGlnMetGlnSer 880

Db 2749 GTTCAAAATATGAAACCAGAGGAGTATGCTCATAGATTTTGAATTACAGATGCAAGT 2808

QY 881 IleMetGluLysLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAsnPro 900

Db 2809 ATAATGGAAAAAGAAATGAAACCAAGAGAAATATTGCCAAGCATTACAAGAATAACCCA 2868

QY 901 SerLeuIleThrPheLeuCysLysAsnCysSerValLeuAlaCysSerGlyGluAspIle 920

Db 2869 TCACATAATACTTTCTTTGCAAAAACCTGCAGTGTGTAGCCTGTTCTGGGGAAGATATC 2928

QY 921 HisValIleGluLysMetHisHisValAsnMetThrProGluPheLysGluLeuTyrIle 940

Db 2929 CATGTAATTGAGAAAAATGCATCACGTCAATATGACCCAGAAATTCAGGAACATTTTACATT 2988

QY 941 ValArgGluAsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluIle 960

Db 2989 GTAAGAGAAAAACAAACCACTGCAAAAGAGTGTGCCGACTATCAAAATAAATGGTGAATC 3048

QY 961 IleCysLysCysGlyGlnAlaTrpGlyThrMetMetValHisLysGlyLeuAspLeuPro 980

Db 3049 ATCTGCAATGTGGCCAGGCTTGGGGAACAATGATGGTGCACAAAGGCTTAGATTGCGCT 3108

QY 981 CysLeuLysIleArgAsnPheValValPheLysAsnAsnSerThrLysLysGlnTyr 1000

Db 3109 TGCTCAAAATAAGGAATTTTGTAGTGGTTTCAAAATAATTCACAAAGAAACAATAC 3168

QY 1001 LysLysTrpValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCysCysLeu 1020

Db 3169 AAAAAAGTGGTAGAATTACCTATCACATTTCCCAATCTTGACTATTCAAGATGCTGTTTA 3228

QY 1021 PheSerAspGluAsp 1025

Db 3229 TTTAGTGTATGAGGAT 3243

RESULT 4

US-10-055-475-8

; Sequence 8, Application US/10055475

; Publication No. US20030022855A1

; GENERAL INFORMATION:

; APPLICANT: Fisher, Paul B.

; APPLICANT: Kang, Dong-Chul

; APPLICANT: Gopalkrishnan, Rahul V.

; TITLE OF INVENTION: USE OF MDA-5 AS AN ANTIVIRAL AND

; TITLE OF INVENTION: ANTIPROLIFERATIVE AGENT

; FILE REFERENCE: A34614-A-PCT-USA-A (070050.1921)

; CURRENT APPLICATION NUMBER: US/10/055,475

; CURRENT FILING DATE: 2002-01-22

; PRIOR APPLICATION NUMBER: PCT/US01/0696Q

; PRIOR FILING DATE: 2001-02-28

; PRIOR APPLICATION NUMBER: 09/515,363

; PRIOR FILING DATE: 2000-02-29

; NUMBER OF SEQ ID NOS: 17

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 8

; LENGTH: 3365

; TYPE: DNA

; ORGANISM: homo sapiens

US-10-055-475-8

Alignment Scores:

Pred. No.:	0	Length:	3365
Score:	5307.00	Matches:	1024
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	99.90%	Mismatches:	0
Query Match:	99.92%	Indels:	0
DB:	14	Gaps:	0

US-09-515-363C-2 (1-1025) x US-10-055-475-8 (1-3365)

QY	1	MetSerAsnGlyTyrSerThrAspGluAsnPheArgTyrLeuIleSerCysPheArgAla	20
DB	169	ATGTCGAATGGGTATTCCACAGACGAGAAATTC	228
QY	21	ArgValLysMetTyrIleGlnValGluProValLeuAspTyrLeuThrPheLeuProAla	40
DB	229	AGGGTGAAAAATGTACATCCAGGTGGAGCCTGTGTCGACTACCTGACCTTTCTGCCTGCA	288
QY	41	GluValLysGluGlnIleGlnArgThrValAlaThrSerGlyAsnMetGlnAlaValGlu	60
DB	289	GAGGTGAAGGAGCAGATTACAGAGACAGTCGCCACCTCCGGGAACATGCAGGCAGTTGAA	348
QY	61	LeuLeuLeuSerThrLeuGluLysGlyValTrpHisLeuGlyTrpThrArgGluPheVal	80
DB	349	CTGCTGCTGAGCACCTTGGAGAAGGGAGTCTGGCACCTTGGTCTGGACTCGGGAATTCGTG	408
QY	81	GluAlaLeuArgArgThrGlySerProLeuAlaAlaArgTyrMetAsnProGluLeuThr	100
DB	409	GAGGCCCTCCGGAGAACCGGCAGCCCTCTGGCCGCCGCTACATGAACCTTGAGCTCAGC	468
QY	101	AspLeuProSerProSerPheGluAsnAlaHisAspGluTyrLeuGlnLeuLeuAsnLeu	120
DB	469	GACTTGCCCTCTCCATCGTTTGAGAACGCTCATGATGAATATCTCCAACCTGCTGAACCTC	528
QY	121	LeuGlnProThrLeuValAspLysLeuLeuValArgAspValLeuAspLysCysMetGlu	140
DB	529	CTTCAGCCCACTCTGGTGGACAAGCTTCTAGTTAGACAGCTCTTGGATAAGTGATGGAG	588
QY	141	GluGluLeuLeuThrIleGluAspArgAsnArgIleAlaAlaGluAsnAsnGlyAsn	160
DB	589	GAGGAACGTGTTGACAATTGAAGACAGAAACCGGATTGCTGCTGCAGAAAACAATGGAAT	648
QY	161	GluSerGlyValArgGluLeuLysArgIleValGlnLysGluAsnTrpPheSerAla	180
DB	649	GAATCAGGTGTAAGAGAGCTACTAAAAGGATTGTCAGAAAGAAAACCTGGTTCTCTGCA	708
QY	181	PheLeuAsnValLeuArgGlnThrGlyAsnAsnGluLeuValGlnGluLeuThrGlySer	200
DB	709	TTTCTGAATGTTCTTCGTCAACACAGGAACAATGAACCTGTCCAAGAGTTAACAGGCTCT	768
QY	201	AspCysSerGluSerAsnAlaGluIleGluAsnLeuSerGlnValAspGlyProGlnVal	220
DB	769	GATTGCTCAGAAAGCAATGCAGAGATTGAGAAATTTATCACAAGTTGATGGTCTCAAGTG	828
QY	221	GluGluGlnLeuLeuSerThrThrValGlnProAsnLeuGluLysGluValTrpGlyMet	240
DB	829	GAAGAGCAACTTCTTTCAACACACAGTTTCAGCCAAATCTGGAGAAGGAGGTCCTGGGCGATG	888
QY	241	GluAsnAsnSerSerGluSerSerPheAlaAspSerSerValValSerGluSerAspThr	260
DB	889	GAGAAATAACTCATCAGAAATCATCTTTTGCAGATTCTTCTGTAGTTTCAGAAATCAGACACA	948
QY	261	SerLeuAlaGluGlySerValSerCysLeuAspGluSerLeuGlyHisAsnSerAsnMet	280
DB	949	AGTTTGGCAGAGGAAGTGTCTAGTCTGCTTTAGATGAAAGTCTTTGGACATAACAGCAACATG	1008
QY	281	GlySerAspSerGlyThrMetGlySerAspSerAspGluGluAsnValAlaAlaArgAla	300
DB	1009	GGCAGTGATTTCAGGCACCATGGGAAGTGATTTCAGATGAAGAGAAATGTGGCAGCAAGAGCA	1068
QY	301	SerProGluProGluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaLeu	320

```
QY 681 PheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGluLys 700
|
|
|
Db 2209 TTTGAAACAATAAAATGTTGAAAGGCTGGCTGAAACCCAGAAATATGAAATGAAAG 2268
|
|
|
QY 701 LeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArgThrGluGluSerAlaArg 720
|
|
|
Db 2269 CTGACCAAAATTAAGAAATACCAATAATGGAGCAATATACTAGGACTGAGGAATCAGCACGA 2328
|
|
|
QY 721 GlyIleIlePheThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTrpIleThrGlu 740
|
|
|
Db 2329 GGAATAATCTTTACAAAAACACGACAGAGTGCATATGCGCTTTCCAGTGGATTACTGAA 2388
|
|
|
QY 741 AsnGluLysPheAlaGluValGlyValLysAlaHisHisLeuIleGlyAlaGlyHisser 760
|
|
|
Db 2389 AATGAAAAATTTGCTGAAGTAGGAGTCAAAGCCCCACCATCTGATTGGAGCTGGACACAGC 2448
|
|
|
QY 761 SerGluPheLysProMetThrGlnAsnGluGlnLysGluValIleSerLysPheArgThr 780
|
|
|
Db 2449 AGTGAGTTCAAAACCCATGACACAGAAATGAACAAAAAGAGTCATTAGTAAATTTGCACT 2508
|
|
|
QY 781 GlyLysIleAsnLeuIleAlaThrThrValAlaGluGluGlyLeuAspIleLysGlu 800
|
|
|
Db 2509 GGAAAAATAAATCTGCTTATCGCTACCACAGTGGCAGAGAAGGTCTGGATATTAAAGAA 2568
|
|
|
QY 801 CysAsnIleValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAlaArg 820
|
|
|
Db 2569 TGTAACATTGTTATCCGTTATGGTCTCGTCACCAATGAATAGCCATGGTCCAGGCCCGT 2628
|
|
|
QY 821 GlyArgAlaArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGlyVal 840
|
|
|
Db 2629 GGTCGAGCCAGAGCTGATGAGAGCACTACGTCCTGGTTGCTCACAGTGGTTCAGGAGTT 2688
|
|
|
QY 841 IleGluHisGluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHisCys 860
|
|
|
Db 2689 ATCGAACATGAGACAGTTAATGATTTCCGAGAGAAGATGATGTATAAAGCTATACATTGT 2748
|
|
|
QY 861 ValGlnAsnMetLysProGluGluTyrAlaHisLysIleLeuGluLeuGlnMetGlnSer 880
|
|
|
Db 2749 GTTCAAAATATGAAACCCAGAGGAGTATGCTCATAAGATTTTGGAAATTACAGATGCAAAAGT 2808
|
|
|
QY 881 IleMetGluLysLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAsnPro 900
|
|
|
Db 2809 ATAATGGAAGAAAGAAATGAAACCAAGAGAAATATTGCCAAGCATTTACAAGATAACCCCA 2868
|
|
|
QY 901 SerLeuIleThrPheLeuCysLysAsnCysSerValLeuAlaCysSerGlyGluAspIle 920
|
|
|
Db 2869 TCACATAATACTTCCTTTGCAAAACTGCAGTGTGCTAGCCTGTTCTGGGGAAGATATC 2928
|
|
|
QY 921 HisValIleGluLysMetHisHisValAsnMetThrProGluPheLysGluLeuTyrIle 940
|
|
|
Db 2929 CATGTAATTGAGAAAAATGCATCACGTCATAATATGACCCCAAGATTCAGGAACITTTACATT 2988
|
|
|
QY 941 ValArgGluAsnLysAlaLeuGlnLysCysAlaAspTyrGlnIleAsnGlyGluIle 960
|
|
|
Db 2989 GTAAGAGAAAAACAAGCACTGCAAAAGAAAGTGTGCCGACTATCAAAATAAATGGTGAATC 3048
|
|
|
QY 961 IleCysLysCysGlyGlnAlaTrpGlyThrMetMetValHisLysGlyLeuAspLeuPro 980
|
|
|
Db 3049 ATCTGCAAAATGTGGCCAGGCTTGGGGAACAATGATGGTGCAAAAGGCTTAGATTGCGCT 3108
|
|
|
QY 981 CysLeuLysIleArgAsnPheValValPheLysAsnAsnSerThrLysLysGlnTyr 1000
|
|
|
Db 3109 TGTCTCAAAATAAGGAATTTTGTAGTGGTTTCAAAAATAAATTCAAACAAAGAAACAAATAC 3168
|
|
|
QY 1001 LysLysTrpValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCysLeu 1020
|
|
|
Db 3169 AAAAAAGTGGGTAGAAATTACCTATCACATTTCCCAATCTTGACTATTTCAGAAATGCTGTGTTA 3228
|
|
|
QY 1021 PheSerAspGluAsp 1025
|
|
|
Db 3229 TTTAGTGATGAGGAT 3243
|
|
|
```

RESULT 5

```
US-10-198-846-13042
; Sequence 13042, Application US/10198846
; Publication No. US20030099974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13042
; LENGTH: 3928
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1, 3925, 3926, 3927, 3928
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-13042
```

Alignment Scores:

```
Pred. No.: 0 Length: 3928
Score: 5303.00 Matches: 1023
Percent Similarity: 99.90% Conservative: 1
Best Local Similarity: 99.80% Mismatches: 1
Query Match: 99.85% Indels: 0
DB: 14 Gaps: 0
```

US-09-515-363C-2 (1-1025) x US-10-198-846-13042 (1-3928)

```
QY 1 MetSerAsnGlyTyrSerThrAspGluAsnPheArgTyrLeuIleSerCysPheArgAla 20
|
|
|
Db 611 ATGTCGAATGGGTATTCCACAGACGAGAAATTCGGCTATCTCATCTCGTGCTTCAGGGCC 670
|
|
|
QY 21 ArgValLysMetTyrIleGlnValGluProValLeuAspTyrLeuThrPheLeuProAla 40
|
|
|
Db 671 AGGTTGAAAATGTACATCCAGGTGGAGCCTGTGCTGGACTACCTGACCTTTCTGCCTGCA 730
|
|
|
QY 41 GluValLysGluGlnIleGlnArgThrValAlaThrSerGlyAsnMetGlnAlaValGlu 60
|
|
|
Db 731 GAGTGAAGGAGCAGATTCAGAGGACAGTCGCCACCTCCGGGAACATGCAGGCAGTTGAA 790
|
|
|
QY 61 LeuLeuLeuSerThrLeuGluLysGlyValTrpHisLeuGlyTrpThrArgGluPheVal 80
|
|
|
Db 791 CTGCTGCTGAGCACCTTTGGAGAAAGGAGTCTGGCACCTTGGTTGGACTCGGGAATTCGTG 850
|
|
|
QY 81 GluAlaLeuArgArgThrGlySerProLeuAlaAlaArgTyrMetAsnProGluLeuThr 100
|
|
|
Db 851 GAGGCCCTCCGGAGAACCGGCAGCCCTCTGGCCGCCCGCTACATGAACCTGAGCTCAGC 910
|
|
|
QY 101 AspLeuProSerProSerPheGluAsnAlaHisAspGluTyrLeuGlnLeuLeuAsnLeu 120
|
|
|
Db 911 GACTTGGCCCTCTCCATCGTTTGAGAACGCTCATGATGAATATCTCCAACCTGCTGAACCTC 970
|
|
|
QY 121 LeuGlnProThrLeuValAspLysLeuLeuValArgAspValLeuAspLysCysMetGlu 140
|
|
|
Db 971 CTTACGCCCACTCTGGTGGACAAGCTTCTAGTTAGAGACGCTCTTTGGATAAGTGCATGGAG 1030
|
|
|
QY 141 GluGluLeuLeuThrIleGluAspArgAsnArgIleAlaAlaGluAsnAsnGlyAsn 160
|
|
|
Db 1031 GAGGAACCTGTTGACAAATTGAAGACAGAAACCGGATTGCTGCTGCAGAAAAACAATGGAAAT 1090
|
|
|
QY 161 GluSerGlyValArgGluLeuLeuLysArgIleValGlnLysGluAsnTrpPheSerAla 180
|
|
|
```


QY	901	SerLeuIleThrPheLeuCysLysAsnCysSerValLeuAlaCysSerGlyGluAspIle	920
DB	3311	TCACATAAATCTTCCTTGTGCAAAAACCTGCAGTGTCTAGCCTGTTCTGGGGAAGATATC	3370
QY	921	HisValIleGluLysMetHisHisValAsnMetThrProGluPheLysGluLeuTyrIle	940
DB	3371	CATGTAATTGAGAAAAATGCATCACGTCAATATGACCCAGAATTCAAGGAACCTTACATT	3430
QY	941	ValArgGluAsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluIle	960
DB	3431	GTAAGAGAAAAACAAACACTGCAAAAGAAGTGTGCCACTATCAATAAATGGTGAAATC	3490
QY	961	IleCysLysCysGlyGlnAlaTrpGlyThrMetMetValHisLysGlyLeuAspLeuPro	980
DB	3491	ATCTGCAAAATGTGGCCAGGCTTGGGGAAACAATGATGGTGACAAAGGCTTAGATTGCCT	3550
QY	981	CysLeuLysIleArgAsnPheValValPheLysAsnAsnSerThrLysLysGlnTyr	1000
DB	3551	TGTCTCAAAATAAGGAAATTTTGTAGTGGTTTTTCAAAAATAAATTCAACAAAGAAACAATAC	3610
QY	1001	LysLysTrpValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCysCysLeu	1020
DB	3611	AAAAAGTGGTAGAATTACCTATCACATTTCCCAATCTTGACTATTCAGAAATGCTGTTA	3670
QY	1021	PheSerAspGluAsp	1025
DB	3671	TTTAGTGATGAGGAT	3685

RESULT 6

```

US-10-055-475-4
; Sequence 4, Application US/10055475
; Publication No. US20030022855A1
; GENERAL INFORMATION:
; APPLICANT: Fisher, Paul B.
; APPLICANT: Kang, Dong-Chul
; APPLICANT: Gopalkrishnan, Rahul V.
; TITLE OF INVENTION: USE OF MDA-5 AS AN ANTIVIRAL AND
; TITLE OF INVENTION: ANTIPROLIFERATIVE AGENT
; FILE REFERENCE: A34614-A-PCT-USA-A (070050.1921)
; CURRENT APPLICATION NUMBER: US/10/055,475
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: PCT/US01/06960
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 09/515,363
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 3627
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-055-475-4

```

Alignment Scores:			
Pred. No.:	0	Length:	3627
Score:	5299.00	Matches:	1023
Percent Similarity:	99.90%	Conservative:	1
Best Local Similarity:	99.80%	Mismatches:	1
Query Match:	99.77%	Indels:	0
DB:	14	Gaps:	0

US-09-515-363C-2 (1-1025) x US-10-055-475-4 (1-3627)

Qy		1	MetSerAsnGlyTyrSerThrAspGluAsnPheArgTyrLeuIleSerCysPheArgAla	20
Dd		169	ATGTCCAATGGGTATTCCACAGACGAGAATTTCGCCTATCTCATCTCGTGCTTCAGGCC	228
Qy		21	ArgVallysMetTyrIleGlnValGluProValLeuAspTyrLeuThrPheLeuProAla	40
Dd		229	AGGGTGAAATGTACATCCAGGTGGAGCCTGTGCTGGACTACCTGACCCTTTCTGCGCTGCA	288
Qy		41	GluVallysGluGlnIleGlnArgThrValAlaThrSerGlyAsnMetGlnAlaValGlu	60

QY 421 LeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPheSerLeuIle 440
Db 1429 CTCTTAAACTTGGAAAATGGAGAGATGCTGGTGTTCATTTGTCCACTTTTCCCTCAT 1488
QY 441 IleIleAspGluCysHisHisThrAsnLysGluAlaValTyrAsnAsnIleMetArgHis 460
Db 1489 ATCAATTGATGAATGTCATCACACCAACAAGAGCAGTGATATAACATCATGAGGCAT 1548
QY 461 TyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysLysGluAsnLysProValIlePro 480
Db 1549 TATTTGATGCAGAAGTTGAAAAACAATAGACTCAAGAAAGAAAAACAACCCAGTGATCCC 1608
QY 481 LeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyGlyAlaThrLysGlnAla 500
Db 1609 CTTCCCTCAGACTGGGACTAACAGCTTCACCTGGTGTGGAGGGGCCACGAAGCAAGCC 1668
QY 501 LysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThrIleLysThr 520
Db 1669 AAAGCTGAAGAACACACATTTTAAAACTATGTGCCAATCTTGTATGCATTTACTATTAAAACT 1728
QY 521 ValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluProCysLysLysPheAla 540
Db 1729 GTTAAAGAAAACCTTGATCAACTGAAAAACCAATAACAGAGCCATGCAAGAAGTTTGCC 1788
QY 541 IleAlaAspAlaThrArgGluAspProPheLysGluLysLeuLeuGluIleMetThrArg 560
Db 1789 ATTGCAAGATGCACCAGAGAAAGATCCATTTTAAAGAGAAACTTCTAGAAATAATGACAAGG 1848
QY 561 IleGlnThrTyrCysGlnMetSerProMetSerAspPheGlyThrGlnProTyrGluGln 580
Db 1849 ATTCAAACTTATTGTCAAAATGATGCCAATGTCCAGATTTTGGAACTCAACCCCTATGAACAA 1908
QY 581 TrpAlaIleGlnMetGluLysLysAlaAlaLysLysGlyAsnArgLysGluArgValCys 600
Db 1909 TGGGCCATTCAAATGGAAAAAAGCTGCAAAAGAGGAATTCGAAAGAACGTGTTGT 1968
QY 601 AlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMetIle 620
Db 1969 GCAGAACATTTGAGGAAGTACAAATGAGGCCCTACAAATTAATGACACAAATTCGAATGATA 2028
QY 621 AspAlaTyrThrHisLeuGluThrPheTyrAsnGluGluLysAspLysLysPheAlaVal 640
Db 2029 GATCGGTATACTCATCTTGAAACTTCTATAATGAAGAGAAAGATAAGAAGTTTCAGTC 2088
QY 641 IleGluAspAspSerAspGluGlyGlyAspAspGluTyrCysAspGlyAspGluAspGlu 660
Db 2089 ATAGAAGATGATAGTGATGAGGGTGGTGATGATGAGTATGTTGTGATGGTGATGAAGATGAG 2148
QY 661 AspAspLeuLysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeuPhe 680
Db 2149 GATGATTTAAAGAAACCTTTGAAACTGGATGAAACAGATAGATTCTCATGACTTTATT 2208
QY 681 PheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGluLys 700
Db 2209 TTTGAAAAACAATAAAATGTTGAAAAGGCTGGCTGAAAACCCAGAAATATGAAAATGAAAAG 2268
QY 701 LeuThrLysLeuArgAsnThrIleMetGluGlnTyrThrArgThrGluGluSerAlaArg 720
Db 2269 CTGACCAAAATTAAGAAATACCATAATGGAGCAATATACTAGGACTGAGGAATCAGCACGA 2328
QY 721 GlyIleIlePheThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTrpIleThrGlu 740
Db 2329 GGAATAATCTTTACAAAAACACGACAGAGTGCATATGCGCTTTCAGTGGATTACTGAA 2388
QY 741 AsnGluLysPheAlaGluValGlyValLysAlaHisLeuIleGlyAlaGlyHisSer 760
Db 2389 AATGAAAAATTTGCTGAAGTAGGAGTCAAAGCCCACCATCTGATTTGGAGCTGGACACAGC 2448
QY 761 SerGluPheLysProMetThrGlnAsnGluGlnLysGluValIleSerLysPheArgThr 780
Db 2449 AGTGAGTTCAAAACCCATGACACAGAAATGAACAAAAAGAGTCATTAGTAAATTTGCGACT 2508

QY 781 GlyLysIleAsnLeuLeuIleAlaThrThrValAlaGluGluGlyLeuAspIleLysGlu 800
Db 2509 GGAAAAATAAATCTGCTTATCGCTACCACAGTGGCAGAAGAAGGTCTGGATATTAAAGAA 2568
QY 801 CysAsnIleValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAlaArg 820
Db 2569 TGTAACATTTGTTATCCGTTTATGCTCTCGTCAACAATGAANAATAGCCATGGTCCAGGCCGT 2628
QY 821 GlyArgAlaArgAlaAspGluSerThrTyrValLeuValAlaAlaHisSerGlySerGlyVal 840
Db 2629 GGTCGAGCCAGAGCTGATGAGAGCACCTACGTCCTGGTTGCTCACAGTGGTTCAGAGATT 2688
QY 841 IleGluHisGluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHisCys 860
Db 2689 ATCGAACGTGAGACAGTTAATGATTTCCGAGAGAAGATGATGTATAAGCTATACATTGT 2748
QY 861 ValGlnAsnMetLysProGluGluTyrAlaHisLysIleLeuGluLeuGlnMetGlnSer 880
Db 2749 GTTCAAAATATGAACCCAGAGGATGATGCTCATAAGATTTTGGAAATTACAGATGCAAGT 2808
QY 881 IleMetGluLysLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAsnPro 900
Db 2809 ATAATGGAAGAAAGAAATGAAAACCAAGAGAAATATTGCCAAGCATTTACAAGAATAAACCCA 2868
QY 901 SerLeuIleThrPheLeuCysLysAsnCysSerValLeuAlaCysSerGlyGluAspIle 920
Db 2869 TCACATAATAACTTTCCTTTGCAAAAAAGTGCAGTGTGTAGCCTGTTCTGGGGAAGATATC 2928
QY 921 HisValIleGluLysMetHisHisValAsnMetThrProGluPheLysGluLeuTyrIle 940
Db 2929 CATGTAATTGAGAAAATGTCATCAGTCAATATGACCCCAAGAAATTCAGGAACTTTACATT 2988
QY 941 ValArgGluAsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluIle 960
Db 2989 GTAAGAGAAAAACAAGCACTGCAAAAAGAAAGTGTGCCGACTATCAAAATAAATGGTGAATC 3048
QY 961 IleCysLysCysGlyGlnAlaTrpGlyThrMetMetValHisLysGlyLeuAspLeuPro 980
Db 3049 ATCTGCAAAATGTGCCAGGCTTGGGGAACAATGATGTTGCACAAAAGGCTTAGATTTGCT 3108
QY 981 CysLeuLysIleArgAsnPheValValPheLysAsnAsnSerThrLysLysGlnTyr 1000
Db 3109 TGTCCTCAAAATAAGGAATTTTGTAGTGGTTTTCAAAAAATAATTCAACAAAGAAACAATAC 3168
QY 1001 LysLysTrpValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCysCysLeu 1020
Db 3169 AAAAAAGTGGTAGAATTACCTATCACATTTTCCCAATCTTGACTATTTCAGAAATGCTGTTA 3228
QY 1021 PheSerAspGluAsp 1025
Db 3229 TTTAGTGTGAGGAT 3243
RESULT 7
US-10-275-822A-1
; Sequence 1, Application US/10275822A
; Publication No. US20040086500A1
; GENERAL INFORMATION:
; APPLICANT: Bahr, Georges
; APPLICANT: Cocude, Cecile
; APPLICANT: Capron, Andre
; TITLE OF INVENTION: RH116 Polypeptide and its Fragments and Polynucleotides Encoding
; TITLE OF INVENTION: said Polypeptides and Therapeutic Uses
; FILE REFERENCE: 017753-170
; CURRENT APPLICATION NUMBER: US/10/275,822A
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: PCT/FR01/01441
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: FR 00/06,030
; PRIOR FILING DATE: 2000-05-11
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3372

```
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (155)..(3229)
US-10-275-822A-1

Alignment Scores:
Pred. No.:      0      Length:      3372
Score:      5285.00      Matches:      1020
Percent Similarity:      99.71%      Conservative:      2
Best Local Similarity:      99.51%      Mismatches:      3
Query Match:      99.51%      Indels:      0
DB:      16      Gaps:      0

US-09-515-363C-2 (1-1025) x US-10-275-822A-1 (1-3372)

QY      1 MetSerAsnGlyTyrSerThrAspGluAsnPheArgTyrLeuIleSerCysPheArgAla 20
Db      155 ATGTCGAATGGGTATTCCACAGACGAGAAATTTCCGCTATCTCATCTCGTGCTTCAGGGCC 214

QY      21 ArgValLysMetTyrIleGlnValGluProValLeuAspTyrLeuThrPheLeuProAla 40
Db      215 AGGGTGAATAATGTACATCCAGGTGGAGCCTGTGCTGGACTACCTGACCTTTCTGCGCTGCA 274

QY      41 GluValLysGluGlnIleGlnArgThrValAlaThrSerGlyAsnMetGlnAlaValGlu 60
Db      275 GAGGTGAAGGAGCAGATTTCAGAGGACAGTCGCCACCTCCGGGAACATGCAGGCAGTTGAA 334

QY      61 LeuLeuLeuSerThrLeuGluLysGlyValTrpHisLeuGlyTrpThrArgGluPheVal 80
Db      335 CTGCTGCTGAGCACCTTGGAGAGGGAGTCTGGCACCTTGGTTGGACTCGGGAATTCGTG 394

QY      81 GluAlaLeuArgArgThrGlySerProLeuAlaAlaArgTyrMetAsnProGluLeuThr 100
Db      395 GAGGCCCTCCGGAGAACCCGGCAGCCCTCTGGCCGCCCGCTACATGAACCCCTGAGCTCAG 454

QY      101 AspLeuProSerProSerPheGluAsnAlaHisAspGluTyrLeuGlnLeuLeuAsnLeu 120
Db      455 GACTTGCCCTCTCCATCGTTTGAAACGCTCATGATGAATATCTCCAACCTGCTGAACCTC 514

QY      121 LeuGlnProThrLeuValAspLysLeuLeuValArgAspValLeuAspLysCysMetGlu 140
Db      515 CTTTCAGCCCACTCTGGTGGACAAGCTTCTAGTTAGAGAGCTCTTGATTAAGTGCATGGAG 574

QY      141 GluGluLeuLeuThrIleGluAspArgAsnArgIleAlaAlaGluAsnAsnGlyAsn 160
Db      575 GAGGAACTGTTGACAAATTGAAGACAGAAACCGGATTGTGCTGCAGAAACAAATGGAAAT 634

QY      161 GluSerGlyValArgGluLeuLeuLysArgIleValGlnLysGluAsnTrpPheSerAla 180
Db      635 GAATCAGGTGTAAGAGAGCTACTAAAAAGGATTGTGCAGAAAGAAACTGGTTCTCTGCA 694

QY      181 PheLeuAsnValLeuArgGlnThrGlyAsnAsnGluLeuValGlnGluLeuThrGlySer 200
Db      695 TTTTCTGAATGTTCTTCGTCAACACAGGAAACAATGAACCTTGCCAAGATTGATGGTCTCAAGTG 754

QY      201 AspCysSerGluSerAsnAlaGluIleGluAsnLeuSerGlnValAspGlyProGlnVal 220
Db      755 GATTGCTCAGAAAGCAATGCAGAGATTGAGAAATTTATCAAGATTGATGGTCTCAAGTG 814

QY      221 GluGluGlnLeuLeuSerThrThrValGlnProAsnLeuGluLysGluValTrpGlyMet 240
Db      815 GAAGAGCAACTTCTTTCAACCACAGTTTCAGCCAAATCTGGAGAAGGAGGTCCTGGGGCATG 874

QY      241 GluAsnAsnSerSerGluSerSerPheAlaAspSerSerValValSerGluSerAspThr 260
Db      875 GAGAATAACTCATCAGAATCATCTTTTGCAGATTCTTCTGTAGTTTCAGAAATCAGACACA 934

QY      261 SerLeuAlaGluGlySerValSerCysLeuAspGluSerLeuGlyHisAsnSerAsnMet 280
Db      935 AGTTTGGCAGAGGAAGTGTTCAGCTGCTTAGATGAAAGTCTTGGACATTAACAGCAACATG 994
```

```
QY      281 GlySerAspSerGlyThrMetGlySerAspSerAspGluGluAsnValAlaAlaArgAla 300
Db      995 GGCAGTGATTTCAGGCACCATGGGAAGTGATTTCAGATGAAGAGATGTGGCAGCAAGAGCA 1054

QY      301 SerProGluProGluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaLeu 320
Db      1055 TCCCCGGAGCCAGAACTCCAGCTCAGGCTTACCAAAATGGAAGTTGCCAGCCAGCCTTG 1114

QY      321 GluGlyLysAsnIleIleCysLeuProThrGlySerGlyLysThrArgValAlaVal 340
Db      1115 GAAGGGAAGAAATATCATCTGCTCCCTACAGGGAGTGGAAAAAACAGAGTGGCTGTT 1174

QY      341 TyrIleAlaLysAspHisLeuAspLysLysLysAlaSerGluProGlyLysValIle 360
Db      1175 TACATTGCCAAGGATCACTTAGACAAGAAAAAAGCATCTGAGCCTGGAAAAAGTTATA 1234

QY      361 ValLeuValAsnLysValLeuLeuValGluGlnLeuPheArgLysGluPheGlnProPhe 380
Db      1235 GTTCTTGTCAATAAGGTACTGTAGTTGAACAGCTCTTCCGCAAGGAGTTCCAACCATTT 1294

QY      381 LeuLysLysTyrTyrArgValIleGlyLeuSerGlyAspThrGlnLeuLysIleSerPhe 400
Db      1295 TTGAAGAAATGGTATCGTGTATTATTGGATTAAAGTGGTATACCCAACTGAAAAATATCATTT 1354

QY      401 ProGluValValLysSerCysAspIleIleSerThrAlaGlnIleLeuGluAsnSer 420
Db      1355 CCAGAAATTGTCAAGTCTCTGTATATTATTATCAGTACAGCTCAAATCCTTGAAAACTCC 1414

QY      421 LeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPheSerLeuIle 440
Db      1415 CTCTTAAACTTGGAAAAATGGAGAAGATGCTGGTGTTCAAATTGTCAGACTTTTCTCTTCATT 1474

QY      441 IleIleAspGluCysHisHisThrAsnLysGluAlaValTyrAsnAsnIleMetArgHis 460
Db      1475 ATCATTTGATGAATGTTCATCACACCAACAAGAAAGCAGTGTATATAACATCATGAGGCAT 1534

QY      461 TyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysLysGluAsnLysProValIlePro 480
Db      1535 TATTTGATGCAGAAAGTTGAAAAACAATAGACTCAAGAAAGAAAAACAAACCAAGTATTCCC 1594

QY      481 LeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyGlyAlaThrLysGlnAla 500
Db      1595 CTTCTCTCAGATACTCGGACTAACAGCTTCACCTGGTGTGGAGGGGCCACGAAGCAAGCC 1654

QY      501 LysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThrIleLysThr 520
Db      1655 AAAGCTGAAGAACACATTTTAAAACTATGTGCCAATCTTGATGCATTTACTATTAAACT 1714

QY      521 VallysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluProCysLysLysPheAla 540
Db      1715 GTTAAAGAAACCTTGTATCAACTGAAACCAATAACAGGAGCCATGCAAGAAGTTTGCC 1774

QY      541 IleAlaAspAlaThrArgGluAspPropheLysGluLysLeuLeuGluIleMetThrArg 560
Db      1775 ATTGCAGATGCAACCCAGAGAAGATCCATTTTAAAGAGAAACTTCTAGAAATAATGACAAGG 1834

QY      561 IleGlnThrTyrCysGlnMetSerProMetSerAspPheGlyThrGlnProTyrGluGln 580
Db      1835 ATTCAAACTTATTGTCAAATGAGTCCAATGTCCAAATGTCCAGACTTTTGGAACTCAACCTTATGAACAA 1894

QY      581 TrpAlaIleGlnMetGluLysLysAlaAlaLysLysGlyAsnArgLysGluArgValCys 600
Db      1895 TGGGCCATTCAAATGGAAAAAAAAGCTGCAAAAGAAAGGAAATCGCAAGAAAAAGTGTGTGT 1954

QY      601 AlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMetIle 620
Db      1955 GCAGAACATTTGAGGAAGTACAATAAGGCCCTTACAAATTAATGACACAATTCGAATGATA 2014

QY      621 AspAlaTyrThrHisLeuGluThrPheTyrAsnGluGluLysAspLysPheAlaVal 640
Db      2015 GATGCGTATACTCATCTTGAAACTTTTCTATAATGAAGAGAAAGATGAAGAGTTTGCAGTC 2074

QY      641 IleGluAspAspSerAspGluGlyGlyAspAspGluTyrCysAspGlyAspGluAspGlu 660
```


QY 314 GluValAlaGlnProAlaLeuGluGlyLysAsnIleIleCysLeuProThrGlySer 333
Db |||||
422 GAAGTTGCCAGCCAGCCTTGAAGGGAAGAATATCATCATCTGCCTCCCTACAGGAGT 481
QY 334 GlyLysThrArgValAlaValTyrIleAlaLysAspHisLeuAspLysLysLysAla 353
Db |||||
482 GGAAAAACCAGAGTGGCTGTTTACATTGCCAAGGATCACTTAGCAAGAAGAAAAAGCA 541
QY 354 SerGluProGlyLysValIleValLeuValAsnLysValLeuValGluGlnLeuPhe 373
Db |||||
542 TCTGAGCCTGGAAAAAGTTATAGTTCTTGTCAATAAGGTACTGCTAGTTGAACAGCTCTTC 601
QY 374 ArgLysGluPheGlnProPheLeuLysLysTyrTyrArgValIleGlyLeuSerGlyAsp 393
Db |||||
602 CGCAAGGAGTTCACACCATTTTGAAGAAATGGTATCGTGTATTGGATTAAGTGGTGAT 661
QY 394 ThrGlnLeuLysIleSerPheProGluValValLysSerCysAspIleIleIleSerThr 413
Db |||||
662 ACCCAACTGAAATAATATCATTTCCAGAAAGTTGTCAAGTCTCTGTGATATTATTATCAGTACA 721
QY 414 AlaGlnIleLeuGluAsnSerLeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGln 433
Db |||||
722 GCTCAATCCTTGAAAACTCCCTCTTAAACTTGGAAAAATGGAGAAGATGCTGGTGTCAA 781
QY 434 LeuSerAspPheSerLeuIleIleIleAspGluCysHisThrAsnLysGluAlaVal 453
Db |||||
782 TTGTCAAGCTTTTCCCTTCATTATCATTTGATGAATGTCAACACCAACAAGACGAGTG 841
QY 454 TyrAsnAsnIleMetArgHisTyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysLys 473
Db |||||
842 TATAATAACATCATGAGGCATTATTTGATGCAGAAGTTGAAAAACAATAGACTCAAGAAA 901
QY 474 GluAsnLysProValIleProLeuProGlnIleLeuGlyLeuThrAlaSerProGlyVal 493
Db |||||
902 GAAACAAACCAGTGATTTCCCTTCNTCAGATACTGGGACTAACAGCTTCACTGGTGTT 961
QY 494 GlyGlyAlaThrLysGlnAlaLysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeu 513
Db |||||
962 GGAGGGGCCACGAAGCAAGCCAAAGCTGAAGAACACATTTTAAAACTATGTGCCAATCTT 1021
QY 514 AspAlaPheThrIleLysThrValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGln 533
Db |||||
1022 GATGCATTTACTATTAAAACTGTTAAAGAAAAACCTTGATCAACTGAAAAACCAATACAG 1081
QY 534 GluProCysLysLysPheAlaIleAlaAspAlaThrArgGluAspPropheLysGluLys 553
Db |||||
1082 GAGCCATGCAAGAAGTTTGCCATTGCAGATGCAACCAGAGAGATCCATTTAAAGAGAAA 1141
QY 554 LeuLeuGluIleMetThrArgIleGlnThrTyrCysGlnMetSerProMetSerAspPhe 573
Db |||||
1142 CTTCTAGAAATAATGACAAGGATTCAAACTTATTGTCAAATGAGTCCAATGTCAGATTTT 1201
QY 574 GlyThrGlnProTyrGluGlnTrpAlaIleGlnMetGluLysLysAlaAlaLysLysGly 593
Db |||||
1202 GGAACCTCAACCCCTATGAACAATGGGCCATTCAAATGGAAAAAAAAGCTGCAAAAGAGGA 1261
QY 594 AsnArgLysGluArgValCysAlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIle 613
Db |||||
1262 AATCGCAAAGAAAGTGTGTCAGAACATTTGAGGAAGTACAATAAGGCCCTACAAATT 1321
QY 614 AsnAspThrIleArgMetIleAspAlaTyrThrHisLeuGluThrPheTyrAsnGluGlu 633
Db |||||
1322 AATGACACAATTGCAATGATAGATGCGTATACTCATCTTGAAACTTTCTATAATGAAGAG 1381
QY 634 LysAspLysLysPheAlaValIleGluAspAspSerAspGluGlyGlyAspAspGluTyr 653
Db |||||
1382 AAAGATAAGAAGTTTGAGTTCATAGAGATGATAGTATGAGGGTGGTATGATGATGATAT 1441

RESULT 11

US-10-641-643-48

; Sequence 48, Application US/10641643

; Publication No. US20040077003A1

; GENERAL INFORMATION:

APPLICANT: Cocks, Benjamin G.
Susan G. Stuart
Jeffrey J. Seilhamer
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL
GENE EXPRESSION

NUMBER OF SEQUENCES: 1508

CORRESPONDENCE ADDRESS:

ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

STREET: 3174 PORTER DRIVE

CITY: PALO ALTO

STATE: CALIFORNIA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/641,643

FILING DATE: 14-Aug-2003

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: <Unknown>

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Zeller, Karen J.

REGISTRATION NUMBER: 37,071

REFERENCE/DOCKET NUMBER: PA-0001 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (650) 855-0555

TELEFAX: (650) 845-4166

INFORMATION FOR SEQ ID NO: 48:

SEQUENCE CHARACTERISTICS:

LENGTH: 1392 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: HMC1NOT01

CLONE: 004700

SEQUENCE DESCRIPTION: SEQ ID NO: 48 :

US-10-641-643-48

Alignment Scores:

Pred. No.:	7.04e-165	Length:	1392
Score:	2072.00	Matches:	400
Percent Similarity:	94.79%	Conservative:	0
Best Local Similarity:	94.79%	Mismatches:	22
Query Match:	39.01%	Indels:	0
DB:	16	Gaps:	0

US-09-515-363C-2 (1-1025) x US-10-641-643-48 (1-1392)

QY	604	LeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMetIleAspAlaTyr	623
Db	3	TTGAGGAAGTACAATGAGGCCCTACAAATTAATGACACAATTCGAATGATAGATGCGTAT	62
QY	624	ThrHisLeuGluThrPheTyrAsnGluGluLysAspLysPheAlaValIleGluAsp	643
Db	63	ACTCATCTTGAAACTTTCTATAATGAAGAGAAAGATAAGAAAGTTTCAGTGCATANNNNNN	122
QY	644	AspSerAspGluGlyGlyAspAspGluTyrCysAspGlyAspGluAspGluAspLeu	663
Db	123	NNNTTA	182
QY	664	LysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeuPhePheGluAsn	683
Db	183	AAGAAACCTTTGAAACTGGATGAAACAGATAGATTTCTCATGACTTTATTTTTTGAAC	242
QY	684	AsnLysMetLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGluLysLeuThrLys	703
Db	243	AATAAAATGTTGAAAAGGCTGGCTGAAAACCCAGAAATATGAAAAGCTGACCAAA	302

QY	704	LeuArgAsnThrIleMetGluGlnTyrThrArgThrGluGluSerAlaArgGlyIle	723
DB	303	TTAAGAAATACCATATATGGAGCAATATACCTAGGACTGAGGAATCAGCACGAGGAATAATC	362
QY	724	PheThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTrpIleThrGluAsnGluLys	743
DB	363	TTTACAAAAACACGACAGAGTGCAATATCGCGCTTTCCAGTGGATTACTGAAATGAAAAA	422
QY	744	PheAlaGluValGlyValLysAlaHisHisLeuIleGlyAlaGlyHisSerSerGluPhe	763
DB	423	TTTGTGAAGTAGGAGTCAAAGCCCAACCATCTGTATTGGAGCTGGACACAGCAGTGAGTTC	482
QY	764	LysProMetThrGlnAsnGluGlnLysGluValIleSerLysPheArgThrGlyLysIle	783
DB	483	AAACCCATGACACAGAAATGAACAAAAAGAAAGTCATTAGTAATTTTCGCACTGGAAAAATA	542
QY	784	AsnLeuLeuIleAlaThrThrValAlaGluGluGlyLeuAspIleLysGluCysAsnIle	803
DB	543	AATCTGCTTATCGCTACACAGTGGCAGAGAGGTCTGGATATTAAAGAAATGTAACATT	602
QY	804	ValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAlaArgGlyArgAla	823
DB	603	GTTATCCGTTATGGTCTCGTCACCAATGAAATAGCCATGGTCCAGGCCCGTGGTCGAGCC	662
QY	824	ArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGlyValIleGluHis	843
DB	663	AGAGCTGATGAGAGCACCTACGTCCTGGTTGCTCACAGTGGTTCAGGAGTTATCGAACGT	722
QY	844	GluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHisCysValGlnAsn	863
DB	723	GAGACAGTTAATGATTTCCGAGAGAGATGATGTATAAAGCTATACATTGTGTCAAAT	782
QY	864	MetLysProGluGluTyrAlaHisLysIleLeuGluLeuGlnMetGlnSerIleMetGlu	883
DB	783	ATGAAACCAGAGGAGTATGCTCATAGATTTTGGAAATTACAGATGCAAGTATAATGGAA	842
QY	884	LysLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAsnProSerLeuIle	903
DB	843	AAGAAAAATGAAACCACAGAGAAATATTGCCAAGCATTACAAGAAATAACCCATCAATA	902
QY	904	ThrPheLeuCysLysAsnCysSerValLeuAlaCysSerGlyGluAspIleHisValIle	923
DB	903	ACTTTTCCTTTGCAAAAACCTGCAGTGTGCTAGCCTGTTCTGGGGAAGATATCCATGTAAT	962
QY	924	GluLysMetHisHisValAsnMetThrProGluPheLysGluLeuTyrIleValArgGlu	943
DB	963	GAGAAAAATGCATCACGTCAATATGACCCAGAAATTCAGGAACTTTACATTGTAAGAGAA	1022
QY	944	AsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluIleIleCysLys	963
DB	1023	AACAAAGCACTGCAAAAGAAAGTGTGCCGACTATCAATAAATGGTGAAATCATCTGCAAA	1082
QY	964	CysGlyGlnAlaTyrGlyThrMetMetValHisLysGlyLeuAspLeuProCysLeuLys	983
DB	1083	TGTGGCCAGGCTTGGGGAAACAATGATGGTGCACAAAGGCTTAGATTTGCCTTGTCTCAAA	1142
QY	984	IleArgAsnPheValValPheLysAsnAsnSerThrLysLysGlnTyrLysLysTrp	1003
DB	1143	ATAAGGAATTTTGTAGTGGTTTCAAAAATAATTTCAACAAAGAAACAATACAAAAAGTGG	1202
QY	1004	ValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCysCysLeuPheSerAsp	1023
DB	1203	GTAGAAATTACCTATCACATTTCCCAATCTTGACTATTACAGAAATGCTGTTTATTAGTCAT	1262
QY	1024	GluAsp	1025
DB	1263	GAGGAT	1268

```

; GENERAL INFORMATION:
; APPLICANT: Bahr, Georges
; APPLICANT: Cocude, Cecile
; APPLICANT: Capron, Andre
; TITLE OF INVENTION: RH116 Polypeptide and its Fragments and Polynucleotides Encoding
; TITLE OF INVENTION: said Polypeptides and Therapeutic Uses
; FILE REFERENCE: 017753-170
; CURRENT APPLICATION NUMBER: US/10/275,822A
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: PCT/FR01/01441
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: FR 00/06,030
; PRIOR FILING DATE: 2000-05-11
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 1284
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1261
; OTHER INFORMATION: n = A,T,C or G
US-10-275-822A-7

Alignment Scores:
Pred. No.: 1.96e-158 1284
Score: 1995.00 378
Percent Similarity: 99.47% Conservative: 0
Best Local Similarity: 99.47% Mismatches: 2
Query Match: 37.56% Indels: 0
DB: 16 Gaps: 0

```

RESULT 12
US-10-275-822A-7
; Sequence 7, Application US/10275822A
; Publication No. US20040086500A1

Db	542	GATGAGAGCACCTACGTCCTGGTTGCTCACAGTGGTTTCAGGAGTTATCGAACGTGAGACA	601
QY	846	ValAsnAspPheArgGluLysMetMetTyrLysAlaIleHisCysValGlnAsnMetLys	865
Db	602	GTTAATGATTTCCGAGAGAAGATGATGATATAAAGCTATACATTGTGTTCAAAATATGAAA	661
QY	866	ProGluGluTyrAlaHisLysIleLeuGluLeuGlnMetGlnSerIleMetGluLysLys	885
Db	662	CCAGAGGAGTATGCTCATAGATTTTGGAAATTACAGATGCAAAAGTATAATGGAAGAAAA	721
QY	886	MetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAsnProSerLeuIleThrPhe	905
Db	722	ATGAAACCAAGAGAAATATTGCCAAGCATTACAAGATTAACCCATCACTAATACTTTC	781
QY	906	LeuCysLysAsnCysSerValLeuAlaCysSerGlyGluAspIleHisValIleGluLys	925
Db	782	CTTTGCAAAACTGCAGTGTGCTAGCCTGTTCTGGGGAAGATATCCATGTAATTGAGAAA	841
QY	926	MetHisHisValAsnMetThrProGluPheLysGluLeuTyrIleValArgGluAsnLys	945
Db	842	ATGCATCACGTCATATATGACCCCGAATTCAAGGAACTTTACATTTGTAAGAGAAAAA	901
QY	946	AlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluIleIleCysLysCysGly	965
Db	902	ACACTGCAAAAGAGTGTCCGACTATCAAAATAAATGGTGAATCATCTGCAAAATGTGC	961
QY	966	GlnAlaTrpGlyThrMetMetValHisLysGlyLeuAspLeuProCysLeuLysIleArg	985
Db	962	CAGGCTTGGGGAACAATGATGGTGCACAAAGCCTTAGATTTGCCTTGCTCAAAATAAGG	1021
QY	986	AsnPheValValPheLysAsnAsnSerThrLysLysGlnTyrLysLysLysTrpValGlu	1005
Db	1022	AATTTGTAGTGGTTTTCAAAAATAATTCAACAAAGAAACATACAAAAAGTGGGTAGAA	1081
QY	1006	LeuProIleThrPheProAsnLeuAspTyrSerGluCysCysLeuPheSerAspGluAsp	1025
Db	1082	TTACCTATCACATTTCCCAATCTTGACTATTTCAGAAATGCTGTATTATTAGTGAGGAT	1141
RESULT 13			
US-10-302-172-771			
; Sequence 771, Application US/10302172			
; Publication No. US20040053250A1			
; GENERAL INFORMATION:			
; APPLICANT: Tang, Y. Tom			
; APPLICANT: Xue, Aidong J.			
; APPLICANT: Drmanac, Radoje T.			
; TITLE OF INVENTION: No. US20040053250A1el Arginine-rich Protein-like Nucleic Acids an			
; TITLE OF INVENTION: Polypeptides			
; FILE REFERENCE: 803_1CNCP			
; CURRENT APPLICATION NUMBER: US/10/302,172			
; CURRENT FILING DATE: 2002-11-21			
; PRIOR APPLICATION NUMBER: US 10/225,251			
; PRIOR FILING DATE: 2002-08-20			
; PRIOR APPLICATION NUMBER: PCT US02/05095			
; PRIOR FILING DATE: 2002-03-05			
; PRIOR APPLICATION NUMBER: US 09/799,451			
; PRIOR FILING DATE: 2001-03-05			
; NUMBER OF SEQ ID NOS: 950			
; SOFTWARE: pt_FL_genes Version 2.0			
; SEQ ID NO 771			
; LENGTH: 2590			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
; FEATURE:			
; NAME/KEY: CDS			
; LOCATION: (190)..(2223)			
US-10-302-172-771			
Alignment Scores:			
Pred. No.:	8.14e-104	Length:	2590
Score:	1352.00	Matches:	292
Percent Similarity:	59.69%	Conservative:	136
Best Local Similarity:	40.73%	Mismatches:	237

Query Match:	25.46%	Indels:	52
DB:	16	Gaps:	12
US-09-515-363C-2 (1-1025) x US-10-302-172-771 (1-2590)			
QY	306	LeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAlaLeuGluGlyLysAsnIle	325
Db	190	ATGGAGCTTCGGTCTCTACCAATGGGAGGTGATCATGCTGCCCTGGAGGGCAAGAAATATC	249
QY	326	IleIleCysLeuProThrGlySerGlyLysThrArgValAlaValTyrIleAlaLysAsp	345
Db	250	ATCATCTGGTGCCACGGGTGCCGGAAGACCCGGGGGCTGCTTATGTGCCAAGCGG	309
QY	346	HisLeuAspLysLysLysAlaSerGluProGlyLysValIleValLeuValAsnLys	365
Db	310	CACCTAGAG-----ACTGTGGATGGAGCCAAAGTGGTTGTATTGGTCAACAGG	357
QY	366	ValLeuLeuValGlnLeuPheArgLysGluPheGlnPropheLeuLysLysTrpTyr	385
Db	358	GTGCACCTGGTGACCCAG--CATGGTGAAGAGTTACGGCGCATGCTGGATGGACGCTGG	414
QY	386	ArgValIleGlyLeuSerGlyAspThrGlnLeuLysIleSerPheProGluValValLys	405
Db	415	ACCGTGACAACCCCTGAGTGGGACATGGGACCACGCTGTGGCTTTGGCCACCTGGCCCGG	474
QY	406	SerCysAspIleIleSerThrAlaGlnIleLeuGluAsnSerLeuLeuAsnLeuGlu	425
Db	475	TGCCATGACCTGCTCATCTGCACAGCAGAGCTTCTGCAGATGGCACTGACCCAGCCCGAG	534
QY	426	AsnGlyGluAspAlaGlyValGlnLeuSerAspPheSerLeuIleIleAspGluCys	445
Db	535	-----GAGGAGGAGCACGCTGGAGCTCACTGTCTTCTCCTGTATCGTGGTGGATGATGC	588
QY	446	HisHisThrAsnLysGluAlaValTyrAsnAsnIleMetArgHisTyrLeuMetGlnLys	465
Db	589	CACCACACGCACAAGGACACCCGCTACAAACGTCATCATGAGCCAGTACCTAGAACTTAA	648
QY	466	LeuLysAsnAsnArgLeuLysLysGluAsnLysProValIleProLeuProGlnIleLeu	485
Db	649	CTCCAGAGGGCACAG-----CCGCTACCCCGAGGTGCTG	681
QY	486	GlyLeuThrAlaSerProGlyValGlyAlaThrLysGlnAlaLysAlaGluGluHis	505
Db	682	GGTCTCACAGCCTCCCAGGCACTGGCGGGCCTCCAAACTCGATGGGGCCATCAACCAC	741
QY	506	IleLeuLysLeuCysAlaAsnLeuAspAlaPheThrIleLysThrValLysGluAsnLeu	525
Db	742	GTCTGCACTCTGTGCCAACTTGGACACGCTGGTGCATCATGTCAACCCAGAACTGCTGC	801
QY	526	AspGlnLeuLysAsnGlnIleGlnGluProCysLysLysPheAlaIleAlaAspAlaThr	545
Db	802	CCCCAGCTGCAGGAGCACAGCCCAACAGCCTTGCAAAACAGTACAACCTCTGCCACAGCGC	861
QY	546	ArgGluAspPropheLysGluLysLeuLeuIleMetThrArgIleGlnThrTyrCys	565
Db	862	AGCCAGGATCCGTTTGGGGACTTGTGAAGAAGCTCATGGACCAATCCATGACCCACTG	921
QY	566	GlnMetSerProMetSer---AspPheGlyThrGlnProTyrGluGlnTrpAlaIleGln	584
Db	922	GAGATGCCTGAGTTGAGCCCGGAAATTTGGACGCAATGTATGAGCAGCAGGTGGTGAAG	981
QY	585	MetGluLysLysAlaAlaLysLysGlyAsnArgLysGluArgValCysAlaGluHisLeu	604
Db	982	CTGAGTGAGGCTGCGGCTTTGGTGGGCTTTCAGGAGCAACGGGTGTATGCGCTTCACTG	1041
QY	605	ArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMetIleAspAlaTyrThr	624
Db	1042	AGGCGCTACAATGACGCGCTGCTCATCCATGACACCCGTCGCGCGCTGGATGCTTGGCT	1101
QY	625	HisLeuGluThrPheTyrAsnGluGluLysAspLysLysPheAlaValIleGluAspAsp	644
Db	1102	GCGCTGCAGGATTTCTATCACAGGGAGCACGTCACTAAACCCAGATCCTG-----	1152

QY 645 SerAspGluGlyGlyAspAspGluTyrCysAspGlyAspGluAspGluAspLeuLys 664
Db 1153 -----TGT----- 1155
QY 665 LysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeuPhePheGluAsnAsn 684
Db 1156 -----GCCGAGCGCGGCTGCTGGCCCTGTTCGATGACCGCAAG 1194
QY 685 LysMetLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGluLysLeuThrLysLeu 704
Db 1195 AATGAGCTGGCCCACTTGGCAACTCATGGCCCA---GAGAATCCAAAACCTGGAGATGCTG 1251
QY 705 ArgAsnThrIleMetGluGlnTyrThrArgThrGluGluSerAlaArgGlyIleIlePhe 724
Db 1252 GAAAAGATCCTGCAAAGGCAGTTCAGT--AGCTCTAACAGCCCTCGGGGTATCATCTTC 1308
QY 725 ThrLysThrArgGlnSerAlaTyrAlaLeuSerGlnTyrIleThrGluAsnGluLysPhe 744
Db 1309 ACCCGCACCCGCCAAAGCGCACACTCCCTCCTGTCTGGCTCCAGCAGCAGCGGCCTG 1368
QY 745 AlaGluValGlyValLysAlaHisHisLeuIleGlyAlaGlyHisSerSerGluPheLys 764
Db 1369 CAGACTGTGGACATCCGGGCCAGCTACTGATTGGGCTGGGAACAGCAGCCAGACACC 1428
QY 765 ProMetThrGlnAsnGluGlnLysGluValIleSerLysPheArgThrGlyLysIleAsn 784
Db 1429 CACATGACCCAGAGGACCAGCAAGAAGTGATCCAGAAGTCCAGAATGGAACCTGAAC 1488
QY 785 LeuLeuIleAlaThrThrValAlaGluGluGlyLeuAspIleLysGluCysAsnIleVal 804
Db 1489 CTTCTGGTGGCCACGAGTGTGGCGGAGGAGGGCTGGACATCCACATTGCAATGTGGTG 1548
QY 805 IleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAlaArgGlyArgAlaArg 824
Db 1549 GTGCGTTATGGGCTCTTGACCAATGAATCTCCATGTTCCAGGCCAGGGCCGTGCCCGG 1608
QY 825 AlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGlyValIleGluHisGlu 844
Db 1609 GCCGATCAGAGTGTATACGCGTTGTAGCAACTGAAGGTAGCCGGAGCTGAAGCGGGAG 1668
QY 845 ThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHisCysValGlnAsnMet 864
Db 1669 CTGATCAACGAGCGCTGGAGACGCTGATGGAGCAGGCAGTGGCTGTGTGCAGAAATG 1728
QY 865 LysProGluGluTyrAlaHisLysIleLeuGluLeuGlnMetGlnSerIleMetGluLys 884
Db 1729 GACCAGGCCGAGTACCAGGCCAAGATCCGGGATCTGCAGCAGCAGCCTTGACCAAGCGG 1788
QY 885 LysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAsnProSerLeuIleThr 904
Db 1789 GCGGCCCAGGACGCCAGCGGGAGAACCCAGCGGCAGCAGTTCCTCAGTGGAGCACGTGCAG 1848
QY 905 PheLeuCysLysAsnCysSerValLeuAlaCysSerGlyGluAspIleHisValIleGlu 924
Db 1849 CTACTCTGCATCAACTGCATGGTGGTGTGGGCCATGGCAGCGACCTGCGGAAGGTGGAG 1908
QY 925 LysMetHisHisValAsnMetThrProGluPheLysGluLeuTyr---IleValArgGlu 943
Db 1909 GGCACCCACCATGTCAATGTGAACCCCAACTTCTCGAACTACTATAATGTCTCCAGGGAT 1968
QY 944 AsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluIleIleCysLys 963
Db 1969 CCTGTGGTCATCAACAAAGTCTTCAAGGACTGGAAGCCTGGGGGTGTCATCAGCTGCAGG 2028
QY 964 ---CysGlyGlnAlaTrpGlyThrMetMetValHisLysGlyLeuAspLeuProCysLeu 982
Db 2029 AACTGTGGGAGGTCTGGGGTCTGCAGATGATCTACAAGTCAGTGAAGCTGCCAGTGCTC 2088
QY 983 LysIleArgAsnPheValValPheLysAsnAsnSerThrLysLysGlnTyrLysLys 1002
Db 2089 AAAGTCGCAGC-----ATGCTGCTGGAGACCCCTCAGGGCGGGTCCAGGCCAAAAG 2142
QY 1003 TrpValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCysCys 1019

Db 2143 TGGTCCCGCGTCCCTTCTCCGTGCCTGACTTTGACTTCCTGCAGCATTGT 2193
RESULT 14
US-10-108-260A-2072
; Sequence 2072, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1e1 full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2072
; LENGTH: 2250
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-108-260A-2072
Alignment Scores:
Pred. No.: 4 26e-85 Length: 2250
Score: 1129.00 Matches: 269
Percent Similarity: 51.33% Conservative: 118
Best Local Similarity: 35.68% Mismatches: 214
Query Match: 21.26% Indels: 153
DB: 16 Gaps: 14
US-09-515-363C-2 (1-1025) x US-10-108-260A-2072 (1-2250)
QY 270 LeuAspGluSerLeuGlyHisAsnSerAsnMetGlySerAspSerGlyThrMetGlySer 289
Db 56 CTGGCTCCTTCCCTCAGTTTCAGTTTCCATCTTGGGTCTCTGAGCTGAGCAGAGTGGCAC 115
QY 290 AspSerAspGluGluAsnValAla---AlaArgAlaSerProGluProGluLeuGlnLeu 308
Db 116 CAGGCTGAGTTAAGTGGGACTGCCCTGGGCAGACCTACCTACTAGACAGAAATGGAGCTT 175
QY 309 ArgProTyrGlnMetGluValAlaGlnProAlaLeuGluGlyLysAsnIleIleCys 328
Db 176 CGTCTCTACCAATGGGAGGTGATCATGCTGCCCTGGAGGGCAAGAATATCATCATCTGG 235
QY 329 LeuProThrGlySerGlyLysThrArgValAlaValTyrIleAlaLysAspHisLeuAsp 348
Db 236 CTGCCACACGGTGCCTGGGAAGACCCCGGGCGGCTGCTTATGTGGCCAAAGCGGCACCTAGAG 295
QY 349 LysLysLysLysAlaSerGluProGlyLysValIleValLeuValAsnLysValLeuLeu 368
Db 296 -----ACTGTGGATGGAGCCAAAGGTGGTTGTATTGGTCAACAGGGTGCACCTG 343
QY 369 ValGluGlnLeuPheArgLysGluPheGlnProPheLeuLysLysTrpTyrArgValIle 388
Db 344 GTGACCCAG---CATGGTGAAGAGATTTCAGGCCCATGCTGGATGGACCGCTGGACCGTGACA 400
QY 389 GlyLeuSerGlyAspThrGlnLeuLysIleSerPheProGluValLysSerCysAsp 408
Db 401 ACCCTGAGTGGGACATGGGACCACGTCGTGCTGGCTTTGGCCACCTGGCCCGG----- 451
QY 409 IleIleIleSerThrAlaGlnIleLeuGluAsnSerLeuLeuAsnGluAsnGlyGlu 428
Db 451 ----- 451
QY 429 AspAlaGlyValGlnLeuSerAspPheSerLeuIleIleAspGluCysHisThr 448
Db 452 -----TGCCACCACACG 463
QY 449 AsnLysGluAlaValTyrAsnAsnIleMetArgHisTyrLeuMetGlnLysLeuLysAsn 468
Db 464 CACAAGGACACCGTCTACACGTCATCATGAGCCAGTACCTAGAACTTAAACTCCAGAGG 523
QY 469 AsnArgLeuLysLysGluAsnLysProValIleProLeuProGlnIleLeuGlyLeuThr 488
:::

Db 524 GCACAG-----CCGCTACCCAGAGTGTGGGTCTCACA 556
Qy 489 AlaSerProGlyValGlyGlyAlaThrLysGlnAlaLysAlaGluGluHisIleLeuLys 508
Db. 557 GCCTCCCCAGGCACTGGCGGGCCTCCAAACTCGATGGGGCCATCAACCACGTCCTGCAG 616
Qy 509 LeuCysAlaAsnLeuAspAlaPheThrIleLysThrValLysGluAsnLeuAspGlnLeu 528
Db 617 CTCTGTGCCAACTGGACACGTTGGTGCATGTCACCCCAGAACTGCTGCCCCAGCTG 676
Qy 529 LysAsnGlnIleGlnGluProCysLysLysPheAlaIleAlaAspAlaThrArgGluAsp 548
Db 677 CAGGAGCACAGCCAAACAGCCTTGCAACACAGTACAACCTCTGCCACAGGCGCAGCCAGAT 736
Qy 549 PropheLysGluLysLeuGluIleMetThrArgIleGlnThrTyrCysGlnMetSer 568
Db 737 CCGTTGGGACTGTGTAAGAAAGCTCATGGACCAATCCATGACCACCTGGAGATGCCT 796
Qy 569 ProMetSer---AspPheGlyThrGlnProTyrGluGlnTrpAlaIleGlnMetGluLys 587
Db 797 GAGTTGAGCCGGAAATTTGGACGCAATGTATGAGCAGCAGTGGTGAAGCTGAGTGAG 856
Qy 588 LysAlaAlaLysLysGlyAsnArgLysGluArgValCysAlaGluHisLeuArgLysTyr 607
Db 857 GCTGCGGCTTTGGCTGGGCTTCAGGAGCAACGGGTGTATGCGTTTCACCTGAGCGCTAC 916
Qy 608 AsnGluAlaLeuGlnIleAsnAspThrIleArgMetIleAspAlaTyrThrHisLeuGlu 627
Db 917 AATGACGCGCTGTCTATCCATGACACCGTCGCGCGCTGGATGCCTTGGCTGCGCTGCAG 976
Qy 628 ThrPheTyrAsnGluGluLysAspLysLysPheAlaValIleGluAspSerAspGlu 647
Db 977 GATTTCATCACAGGGAGCACGCTCACTAAACCCAGATCCTG----- 1018
Qy 648 GlyGlyAspAspGluTyrCysAspGlyAspGluAspGluAspLeuLysLysProLeu 667
Db 1019 -----TGT----- 1021
Qy 668 LysLeuAspGluThrAspArgPheLeuMetThrLeuPhePheGluAsnAsnLysMetLeu 687
Db 1022 -----GCCGAGCGCGGCTGCTGGCCCTGTTCCGATGACCGCAAGATGAGCTG 1069
Qy 688 LysArgLeuAlaGluAsnProGluTyrGluAsnGluLysLeuThrLysLeuArgAsnThr 707
Db 1070 GCCCACTTGGSCAACTCATGGCCCA---GAGAATCCAAAACTGGAGATGCTGGAAGATC 1126
Qy 708 IleMetGluGlnTyrThrArgThrGluGluSerAlaArgGlyIleIlePheThrLysThr 727
Db 1127 CTGCAAAGGCAGTTCAGT---AGCTCTAACAGCCCTCGGGGTATCATCTTACCCCGCAC 1183
Qy 728 ArgGlnSerAlaTyrAlaLeuSerGlnTrpIleThrGluAsnGluLysPheAlaGluVal 747
Db 1184 CGCCAAAGCGCACACTCCCTCTGCTCTGCTCCAGCAGCAGCGGCTGCAGACTGTG 1243
Qy 748 GlyValLysAlaHisHisLeuIleGlyAlaGlyHisSerSerGluPheLysProMetThr 767
Db 1244 GACATCCGGGCCAGCTACTGATTGGGCTGGGAACAGCAGCCAGAGCACCCACATGACC 1303
Qy 768 GlnAsnGluGlnLysGluValIleSerLysPheArgThrGlyLysIleAsnLeuLeuIle 787
Db 1304 CAGAGGGACCAGCAAGAAGTATCCAGAAGTTCCAAGATGGAACTTGAACCTTCTGGTG 1363
Qy 788 AlaThrThrValAlaGluGluGlyLeuAspIleLysGluCysAsnIleValIleArgTyr 807
Db 1364 GCCACGAGTGTGGCGGAGGAGGGGCTGGACATCCCACATTGCAATGTGGTGGCTGTTAT 1423
Qy 808 GlyLeuValThrAsnGluIleAlaMetValGlnAlaArgGlyArgAlaArgAlaAspGlu 827
Db 1424 GGGCTCTTGACCAATGAAATCTCCATGGTCCAGGCCAGGGGCGCTGCCCGCGCATCAG 1483
Qy 828 SerThrTyrValLeuValAlaHisSerGlySerGlyValIleGluHisGluThrValAsn 847
Db 1484 AGTGATACGCGTTTGTAGCAACTGAAGGTAGCCGGGAGCTGAAGCGGGAGCTGATCAAC 1543

Qy 848 AspPheArgGluLysMetMetTyrLysAlaIleHisCysValGlnAsnMetLysProGlu 867
Db 1544 GAGCGCGTGGAGACGCTGATGGAGCAGGCACTGGCTGCTGTGCAGAAATGGACCGCC 1603
Qy 868 GluTyrAlaHisLysIleLeuGluLeuGlnMetGlnSerIleMetGluLysLysMetLys 887
Db 1604 GAGTACCAG----- 1612
Qy 888 ThrLysArgAsnIleAlaLys-HisTyrLysAsnAsnProSerLeuIleThrPheLeuCy 907
Db 1613 -----GCCAAGGAATACTATAAT----- 1631
Qy 907 sLysAsnCysSerValLeuAlaCysSerGlyGluAspIleHisValIleGluLysMetH1 927
Db 1631 ----- 1631
Qy 927 sHisValAsnMetThrProGluPheLysGluLeuTyrIleValArgGluAsnLysAlaLe 947
Db 1632 -----GTCTCCAGGGATCCTGTGTGCAT 1654
Qy 947 uGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluIleIleCysLys---CysGlyGl 966
Db 1655 CAACAAAGTCTTCAAGGACTGGAAGCCCTGGGGTGTCTATCAGCTGCAGAACTGTGGGA 1714
Qy 966 nAlaTrpGlyThrMetMetValHisLysGlyLeuAspLeuProCysLeuLysIleArgAs 986
Db 1715 GGTCTGGGCTGCAGATGATCTACAAGTCAGTGAAGTGCAGTGCAGTGCAGTGCAGTGCAG 1774
Qy 986 nPheValValPheLysAsnAsnSerThrLysLysGlnTyrLysLysTrpValGluLe 1006
Db 1775 C-----ATGCTGTGGAGACCCCTCAGGGCGGATCCAGGCCAAAGTGGTCCCGGT 1828
Qy 1006 uProIleThrPheProAsnLeuAspTyrSerGluCysCys 1019
Db 1829 GCCCTTCTCCGTGCCTGACTTTGACTTCTCTGCAGCATTGT 1868

RESULT 15

US-10-172-118-1472
; Sequence 1472, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1472
; LENGTH: 3065
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NM_014314
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-1472

Alignment Scores:
Pred. No.: 1.14e-81 Length: 3065
Score: 1090.50 Matches: 327
Percent Similarity: 47.33% Conservative: 169
Best Local Similarity: 31.20% Mismatches: 365
Query Match: 20.53% Indels: 187
DB: 15 Gaps: 36

```

US-09-515-363C-2 (1-1025) x US-10-172-118-1472 (1-3065)
QY 4 GlyTyrSerThrAspGluAsnPheArgTyrLeuIleSerCysPheArgAlaArgVallys 23
Db 155 GGCATGACCACCGAGCAG-----CGACGCAGCCCTGCAAGCCTTCCAGGATTATATCCGG 208
QY 24 MetTyrIleGlnValGluProValLeuAspTyrLeuThr-----PheLeuProAlaGlu 41
Db 209 AAGACCCTGGACCCCTACCTACATCTGAGCTACATGGCCCCCTGGTTTAGGGAGGAAGAG 268
QY 42 VallysGluGlnIleGlnArgThrValAlaThrSerGlyAsnMetGlnAlaValGluLeu 61
Db 269 GTG---CAGTATATTTCAGGCTGAGAAAAACAACAAGGCCCAATGGAGGCTGCCACACTT 325
QY 62 LeuLeuSerThrLeuGluLysGlyValTrpHisLeuGlyTrpThrArgGluPheValGlu 81
Db 326 TTTCTCAAGTTCCTGTTGGAG---CTCCAGGAGGAAGGCTGGTTCCTGGCTTTTGGAT 382
QY 82 AlaLeuArgArgThrGlySerProLeuAlaAlaArgTyrMetAsnProGluLeuThrAsp 101
Db 383 GCCCTAGACCATGCAGGT-----TATTCTGGACTTTATGAAGCCATTGAAAGTTGGAT 436
QY 102 LeuProSerProSerPheGluAsnAlaHis-----AspGluTyrLeuGlnLeuLeuAsn 119
Db 437 -----TTCAAAAAAATTGAAAAGTTGGAGGAGTATAGATTACTTTAAAA 481
QY 120 LeuLeuGlnProThrLeuValAspLysLeuLeuValArgAspValLeuAspLysCysMet 139
Db 482 CGTTTACAACCCAGAATTTAAAACCCAGAATTATCCCAACCGATATCATTTCTGAT---CTG 538
QY 140 GluGluGluLeuLeuThrIleGluAspArgAsnArgIleAlaAlaGluAsnAsnGly 159
Db 539 TCTGAATGTTTAATTAATCAGGAATGTGAAGAAATTTCTACAGATTGCTCTACTAAGGGG 598
QY 160 AsnGluSerGlyValArgGluLeuLysArgIleValGln-----LysGluAsnTrp 177
Db 599 ATGATGGCAGGTGCAGAGAAAATTGGTGAATGCCTTCTCAGATCAGACAAGGAAAACCTGG 658
QY 178 PheSerAlaPheLeuAsnValLeuArgGlnThrGlyAsnAsnGluLeuValGlnGluLeu 197
Db 659 CCCAAAACCTTG-----AAACTTGCTTTGGAGAAA 688
QY 198 ThrGlySerAspCysSerGluSerAsnAlaGluIleGluAsnLeuSerGlnValAspGly 217
Db 689 GAAAGGAACAAGTTTCAGTGAACCTGTGGATTGTAGAGAAAGGTATAAAAGATGTTGAAACA 748
QY 218 ProGlnValGluGlnLeuLeuSerThrThrValGlnProAsnLeuGluLysGluVal 237
Db 749 GAAGATCTTGAGGATAAGATGGAACCTTCTGACATACAG-----ATT 790
QY 238 TrpGlyMetGluAsn-----AsnSerSerGluSerSerPheAlaAspSerSer 253
Db 791 TTCTACCAAGAAGATCCAGAATGCCAGAATCTTAGTGAGAATTTCATGTCACCTTCAGAA 850
QY 254 ValValSerGluSerAspThrSerLeuAlaGluGlySerValSerCysLeuAspGluSer 273
Db 851 GTG-----TCTGATACAAACTTG----- 868
QY 274 LeuGlyHisAsnSerAsnMetGlySerAspSerGlyThrMetGlySerAspSerAspGlu 293
Db 868 ----- 868
QY 294 GluAsnValAlaAlaArgAlaSerProGluProGluLeuGlnLeuArgProTyrGlnMet 313
Db 869 -----TACAGCCCA-----TTTAAACCAAGAAATTACCAATTA 901
QY 314 GluValAlaGlnProAlaLeuGluGlyLysAsnIleIleCysLeuProThrGlySer 333
Db 902 GAGCTTGCTTTGCCTGCTATGAAAGGAAAAACAATAATATGTGCTCCTACAGTTGT 961
QY 334 GlyLysThrArgValAlaValTyrIleAlaLysAspHisLeuAspLysLysLysAla 353

```

```

962 GGAAAAAACCTTTGTTTCACTGCTTATATGTGAACATCATCTT---AAAAAATTCCCAAA 1018
354 SerGluProGlyLysValIleValLeuValAsnLysValLeuLeuValGluGln----- 371
1019 GGACAAAAGGGAAAGTTGTCCTTTTTCGAATCAGATCCAGTGTATGAACACAGAGAAA 1078
372 ---LeuPheArgLysGluPheGlnProPheLeuLysLysTrpTyrArgValIleGlyLeu 390
1079 TCTGTATTCTCAAAATACTTTTGAA-----AGACATGGGTATAGATTACAGGCATT 1129
391 SerGlyAspThrGlnLeuLysIleSerPheProGluValValLysSerCysAspIleIle 410
1130 TCTGGAGCAACAGCTGAGAATGTCCAGTGGAAACAGATTGTTGAGAACAAATGACATCATC 1189
411 IleSerThrAlaGlnIleLeuGluAsnSerLeuLeuAsnGluAsnGlyGluAspAla 430
1190 ATTTTAACTCCACAGATTCTTGTGAAC-----AACCTTAAAAAGGGAACGATTCCA 1240
431 GlyValGlnLeuSerAspPheSerLeuIleIleAspGluCysHisHisThrAsnLys 450
1241 -----TCACTATCCATCTTTACTTTTGATGATATTTGATGAATGCCACAACACTAGTAAA 1294
451 GluAlaValTyrAsnAsnIleMetArgHisTyrLeuMetGlnLysLeuLysAsnAsnArg 470
1295 CAACACCCGTACATATGATCATGTTTAATTATCTAGATCAGAAAACCTTGAGGATCTTCA 1354
471 LeuLysLysGluAsnLysProValIleProLeuProGlnIleLeuGlyLeuThrAlaSer 490
1355 -----GGCCCACTGCCCCAGGTCTATGGCTGACTGCCTCG 1390
491 ProGlyValGlyGlyAlaThrLysGlnAlaLysAlaGluGluHisIleLeuLysLeuCys 510
1391 GTTGGTGTGGGGATGCCAAAAACACAGATGAAGCCTTGGATTATATCTGCAAGCTGTGT 1450
511 AlaAsnLeuAspAlaPheThrIleLysThrValLysGluAsnLeuAspGlnLeuLysAsn 530
1451 GCTTCTCTTGATGCGTCAGTGATAGCAACAGTCAACACAACTCGAGGAACCTGGAGCAA 1510
531 GlnIleGlnGluProCysLysLysPheAlaIleAlaAspAlaThrArgGluAspProphe 550
1511 GTTGTTTATTAAGCCCCAGAAAGTTTTCAGGAAAGTGAATCACCGATTAGCGACAAATTT 1570
551 LysGlu-----Lys 553
1571 AAATACATCATAGCTCAGCTGATGAGGGACACAGAGAGTCTGGCAAGAGAAATCTGCAAA 1630
554 LeuLeuGluIleMetThrArgIleGlnThrTyrCysGlnMetSerProMetSerAspPhe 573
1631 GACCTCGAAAACCTTATCTCAAATTTCAAAAT-----AGGGAATTT 1669
574 GlyThrGlnProTyrGluGlnTrpAlaIleGlnMetGluLysLysAla----- 589
1670 GGAACACAGAAATATGAACAATGGATTGTTACAGTTTCAGAAAGCATGTCATGTTCCAG 1729
590 AlaLysLysGlyAsnArgLysGluArgValCys-----AlaGluHis 603
1730 ATGCCAGACAAAGATGAAGAGACAGGAGATTGTTAAAGCCCTGTTTTTATACACTTCACAT 1789
604 LeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMetIleAspAlaTyr 623
1790 TTGCGGAAATATAATGATGCCCTCATTTATCAGTGAGCATGCACGAATGAAGATGCTCTG 1849
624 ThrHisLeuGluThrPheTyrAsnGluGluLysAspLysLysPheAlaValIleGluAsp 643
1850 GATTACTTGAAGACTTCTTCAGCAATGTCCGACGACGAGGATTTCGAACAGATTGAGCAA 1909
644 AspSerAspGluGlyGlyAspAspGluTyrCysAspGlyAspGluAspAspLeu 663
1910 GAT----- 1912
664 LysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeuPhePheGluAsn 683
1913 -----CTTACTCAGAGATTGAAAGAAAAG 1936

```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 26, 2005, 19:36:55 ; Search time 49 Seconds
(without alignments)
2012.698 Million cell updates/sec

Title: US-09-515-363C-2
Perfect score: 5311
Sequence: 1 MSGYSTDENFRYLISCFRA.....LPITFPNLDYSECCLFSDDED 1025

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: Pirl:*
2: Pirl2:*
3: Pirl3:*
4: Pirl4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	687	12.9	1037	2 T32534	hypothetical prote
2	654	12.3	811	2 T30968	helicase homolog C
3	642.5	12.1	956	2 A88708	protein C01B10.1 [
4	516	9.7	398	2 T46312	hypothetical prote
5	499	9.4	1134	2 T20332	hypothetical prote
6	453	8.5	752	2 D75219	ATP-dependent RNA
7	441.5	8.3	650	2 D71203	probable ATP-depen
8	417	7.9	741	1 A69432	ATP-dependent RNA
9	402.5	7.6	1374	2 S62524	probable RNA helic
10	392.5	7.4	784	2 H64487	eIF-4A family prob
11	350	6.6	1822	2 S44849	K12H4.8 protein -
12	342	6.4	738	1 E69055	ATP-dependent RNA
13	313.5	5.9	784	2 D84386	ATP-dependent RNA
14	303.5	5.7	993	2 S48436	probable RNA helic
15	284	5.3	502	2 T31323	ATP-dependent RNA
16	278.5	5.2	839	2 T39190	probable ATP-depen
17	223	4.2	557	2 G69958	SNF2 helicase homo
18	218	4.1	1979	2 C71622	hypothetical prote
19	217.5	4.1	2663	1 S28261	centromere protein
20	214.5	4.0	1790	2 S67593	transport protein
21	209.5	3.9	449	2 C71156	probable helicase
22	205	3.9	453	2 B75006	DNA repair protein
23	205	3.9	967	2 D72308	conserved hypothet
24	204.5	3.9	2163	2 S50675	pre-mRNA splicing
25	204	3.8	1939	2 T18372	repeat organellar
26	201.5	3.8	1063	2 T38420	probable DNA helic
27	198.5	3.7	2253	2 T30336	nuclear/mitotic ap
28	197.5	3.7	1163	2 G97236	ATPase involved in
29	195	3.7	1531	2 T48946	hypothetical prote

30	195	3.7	3259	1 A56539	giantin - human
31	194.5	3.7	986	2 E90596	restriction-modifi
32	194.5	3.7	1256	2 T26101	hypothetical prote
33	194	3.7	1039	2 E72734	hypothetical prote
34	193.5	3.6	663	2 G89848	exinuclease ABC su
35	193.5	3.6	2017	1 A36014	myosin heavy chain
36	193.5	3.6	2057	2 S61477	myosin II heavy ch
37	193.5	3.6	3225	2 I52300	giantin - human
38	193	3.6	715	2 B75135	DNA helicase relat
39	193	3.6	986	2 H90565	restriction modifi
40	192.5	3.6	718	2 E71100	hypothetical prote
41	192.5	3.6	1300	2 I53799	CG1 protein - huma
42	192	3.6	986	2 S49394	HsdR1 protein - My
43	192	3.6	1356	2 S32763	kinectin 1 - human
44	191.5	3.6	943	2 F69543	ATP-dependent RNA
45	191	3.6	1922	2 T00637	hypothetical prote

ALIGNMENTS

RESULT 1

T32534
hypothetical protein F15B10.2 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C;Accession: T32534
R;Wohldmann, P.; Murray, J.
submitted to the EMBL Data Library, December 1997
A;Description: The sequence of C. elegans cosmid F15B10.
A;Reference number: Z21186
A;Accession: T32534
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-1037 <WOH>
A;Cross-references: UNIPROT:O44165; EMBL:AF036696; PIDN:AAB88350.1; GSPDB:GN00022; CESP
A;Experimental source: strain Bristol N2; clone F15B10
C;Genetics:
A;Gene: CESP:F15B10.2
A;Map position: 4
A;Introns: 30/3; 89/1; 111/2; 218/1; 258/2; 288/1; 343/3; 370/3; 400/2; 443/1; 520/1; 5

Query Match	12.9%	Score	687;	DB	2;	Length	1037;
Best Local Similarity	25.0%	Pred. No.	2.9e-27;				
Matches	251;	Conservative	169;	Mismatches	408;	Indels	174;
Gaps	35;						
Qy	101	DLPPSPFENAHDE--YLQLLNLLQPTLVVDKLLVRDVLDKMEEE-----	142				
Db	63	DLKNSILSNADDERLYKDIMTYLQ-TYLPKCTVHKLLN-CSNREVKLSDFHYLDHFEGF	120				
Qy	143	LLTIEDRNRIAAENNGNESGVRELLKRIVQK---ENWFSAFLN--VLRQTG--NNELVQ	195				
Db	121	LRFIEPKVVLAYLDSYPQYIDAVAVLRKEIERNEEDNQSDFIKKLILRTVPLLGEQAVY	180				
Qy	196	EL--TGSDCSESNAEINLSQVDPQVVEEQLLSTTVQPNLEKEVWGMEN-----N	243				
Db	181	DIMYTISEKSSNNLDVE-----AKQFIA-----KVLRLKNDGFLRFYQIN	221				
Qy	244	SSESSFADSSVWVSEDTSLAEGSVSCLDESIGHNS-----NMGSDSGTMGSDSDE---ENV	296				
Db	222	ASRRQLNGRIYICPVHESATENMVYLGTAAALNTRNRYMINIRVDNIVQENSTPRLVIESV	281				
Qy	297	AARASPEPELQRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKDHLDKKKASEP	356				
Db	282	RQRIHRQRLCLRNQEEELCVALQGKNTIVTAPTSGKTVIAANIIKEHFESRSSEGKR	341				
Qy	357	GKVIVLV-NKVLLVEQLFRKEFPFLKKWYRVIGLSGDTQLKISFPEVVKSCDIIISTAQ	415				
Db	342	FKALFMTFNSMILNQ--AASISSYLDHVYHTQIIQSGD--NVPTRNVIQSKDLIVATPQ	397				
Qy	416	ILENSLLNLENGEDAGVQ-----LSDFSLIIDECHTNKEAVYNNIMRHYLMQKLKNR	470				
Db	398	MIVNLCNEHRNSLDDDESRLDQFLSTFTIIFFECHTNTVKNSPYSNIMREY--HYLKNMG	455				

Db 21 LESIYRDPKGEBCFSELLPLGKIDELKTHS-ENSOEFSKQLYQDLKSNLSIADP-----E 75

QY 129 LLVRDVL-----DKMEEELLTIEDNRNRIAAENNGESGVRELLKRIVQKENWFSAF 181

Db 76 RLYKDIMAYLQSNFPKCILHDVLNYSGRN-VKLSEFRYILDHLEGLRLIEPK-----VV 129

QY 182 LNVLRQTGN-NELVQELTGSDCSESNABIEIENLSQVDPQVBEQQLSTTVQPNLEKEVWGM 240

Db 130 LNYLDSYPQYSEAVKLL--KDEIDRHEEDNN-----DPEFIKKLILRTVPLLGEOAVYDI 182

QY 241 ENNSESSFADSSVSESDTSLA-EGSVSCLDESIGHNSNMGSDSGTMGSDSDENVAAR 299

Db 183 MYTIWEKAQMTSIIINIRVDSIAPENSASRL-----VIESVRQR 221

QY 300 ASPEPELQRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKDLHDKKKASEPGKV 359

Db 222 IHQRQFCLRNQOEELCQVALQGNKTIVTAPTGSKTIVAAIIEKEHFESR---SSEG-- 276

QY 360 IVLVNKVLLVEQLFRKEFPFLKKYRVIGLSGDTQLKISFEFVVVKSCDIIISTAQILEN 419

Db 277 -----KRFKQTQIGSD--NVPTRNVIQSKDLIVATQMIVN 311

QY 420 SLLNLENGEDAGVQ-----LSDFSLLIIDECHHTNKAEVYNNIMRHYLMQKLKNNRLK 472

Db 312 -LCN-EHRDPLDDEYPPPEQFFLSTFTIIFDECHNTVKNSPYSNVVMREY--HYLKNMGNM 367

QY 473 KENKPVIPLPQILGLTASPGVGATQAKAEHILKLCANLDAFTIKTVKENLDQKNQI 532

Db 368 PEGH---SFPQIIGLTASLTGDKKNCMQVRSYIAGLCANMDVKELSIKONLEELLHDN 424

QY 533 QEPCKFAIADATREDP-----FKBKLLIIMTRIQTICYQMSPMDFGTQP-----Y 578

Db 425 PFVTDQVSFCERSNDGPIEMFTKRLKQMMQVEVDLIRTTLKNEPTVKYEIPPTDKEHNR 484

QY 579 EQWAIQMEKKAAGKGNRKERVCAEHLRKYNEALQINDTIRMIDAYTHLETIFYNEEKDKKF 638

Db 485 ENWISQNRNCVSLAGSRNKTLLIEVL-----DVLKVQYMFQSKMIFWNYFRKYRF 534

QY 639 AV--IEDDSDEGGDEYCDGDEDEDDKKPLKLD-----ETDRFLMTLFFENNMMLK 688

Db 535 GFKRIFQFQDCFYALSYNINFNPVALKKYLEKELGPERIRNFTDN--MNLWDN--CHR 590

QY 689 RLAENPEYENEKLTKLRTIMEQYTRTEESARGIIFTKTROSAYALSQWITENEKEFAEVG 748

Db 591 ELVGIGSAENPMIARTVQFILDQNEQTSDFRAIIFVRTKKEADFLNYVL--NDRLHEL 647

QY 749 VKAHLIG---AGHSSEFKPMTQNEQKEVISKFRGTGKINLLIATTVAEGLDIKECNIVI 805

Db 648 IKSDWMSGQKKSASSADISASKQKQMEKLMFADGENQILVSTSVAEGLDIPECSLVI 707

QY 806 RYGLVTNEIAMVQARGARADESTYVLVAHSGSGVIEH--ETVNDFREKMYKAHICVQN 863

Db 708 KINYATNETAHVQRGRARARNKSCVLITNS---IALHVQESNNLAKENLMTETISLIQN 764

QY 864 MKPE-----EYAHKILELQMQSIMEKMKTKRNIACHYKNNPSLITFLCKNCSVLACSG 917

Db 765 SPGEFRQCVDEESNKVWPRIQREDTDKAQRIKEQI-----NRNIVYKIVCMKCDTVLCTN 819

QY 918 EDIHVIEKMHVNMTPPEF-----KELYIVRENKALQKCADYQINGEIIIC---KCGQAWGT 970

Db 820 KDIR-SKNTQYIVCNPGFWSLVRRIPLPQERASNK----FNSTGSIECLGERCGSKLGQ 874

QY 971 MMVHKGLDLPCLKIRNFVVFVKNNSTKKQYKKWVEL-----PITFPNLD 1014

Db 875 LIDVNTVNLPCLVKKSILLIESTNERILVKQWKNILDEHFTPTLTKQRD 924

RESULT 4

T46312

hypothetical protein DKF2p434J1111.1 - human (fragment)

C;Species: Homo sapiens (man)

C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004

C;Accession: T46312

R;Duesterhoeft, A.; Lauber, J.; Mewes, H.W.; Weil, B.; Wiemann, S. submitted to the Protein Sequence Database, January 2000

A;Reference number: Z23036

A;Accession: T46312

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-398 <AAA>

A;Cross-references: UNIPROT:Q9NT04; EMBL:AL137608

A;Experimental source: adult testis; clone DKF2p434J1111

C;Genetics:

A;Note: DKF2p434J1111.1

Query Match 9.7%; Score 516; DB 2; Length 398;

Best Local Similarity 33.3%; Pred. No. 3.7e-19;

Matches 141; Conservative 68; Mismatches 157; Indels 58; Gaps 14;

QY 594 NRKERV-----AEHLRKYNEALQINDTIRMIDAYTHLETIFYNEEKDKKFAVIEDDSDE 647

Db 2 DEESRICKALFLYTSHLRKYNDALIISEHARMKDALDYLKDFFSNVRAAGF----- 52

QY 648 GGDDEYCDGDEDEDDKKPLKLDLTDRLMTLFFENNMMLKRLAENPEYENEKLTKLRT 707

Db 53 -----DEIEQDL-----TQR-----FEELQLELESVSRDPSNENPKLEDLCFI 90

QY 708 IMEQYTRTEESARGIIFTKTROSAYALSQWITENEK--FAEVGVKAHHLIGAGHSSEFKP 765

Db 91 LQEEYHLNPETIT-ILFVKTRALVDALKNWIEGNPKLSFLKPGI-----LTGRGKTNQNTG 145

QY 766 MTQNEQKEVISKFR-TGKINLLIATTVAEGLDIKECNIVIRYGLVTNEIAMVQARGRAR 824

Db 146 MTLPAQKCILDAFKASGDHNLITATSVADEGIDIAQCNLVILYEVGVNVIKMIQTRGRGR 205

QY 825 ADESTYVLVAHSGSGVIEHETVNDFREKMYKAHICVQNMKPEEYAHKILELQMQSIMEK 884

Db 206 ARGSKCFLLT-SNAGVIEKEQINMYKEKMMNDSILRLQTWDEAVFREKILHIQTH---EK 261

QY 885 KMKTKRNIACHY--KNPNSLITFLCKNCSVLACSGEDIHVIKMHVNMTPPEFKELYIVR 942

Db 262 FIRDSQEKPKVPDPKENKKL---LCRKCKALACYTADVRIECHYTVLGDAFKECFVSR 318

QY 943 ENKALQKKCADYQINGEIIIC---KCGQAWGTMVTHKGLDLPCLKIRNFVVFVKNNSTKKQ 999

Db 319 PHPK-PKQFSSFKEKRAKIFCARQNCSDHWGIHVYKTKTEIPVIKIESFVVEDIATGVQTL 377

QY 1000 YKKW 1003

Db 378 YSKW 381

RESULT 5

T20332

hypothetical protein D2005.5 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C;Accession: T20332

R;Wilkinson, J.

submitted to the EMBL Data Library, September 1996

A;Reference number: Z19258

A;Accession: T20332

A;Status: preliminary; translated from GB/EMBL/DBDJ

A;Molecule type: DNA

A;Residues: 1-1134 <WIL>

A;Cross-references: UNIPROT:Q93413; EMBL:Z79752; PIDN:CAB02082.1; GSPDB:GN00019; CESP:D2005

A;Experimental source: clone D2005

C;Genetics:

A;Gene: CESP:D2005.5

A;Map position: 1

A;Introns: 33/2; 91/2; 116/3; 144/1; 319/1; 361/1; 421/3; 482/2; 591/1; 684/3; 757/1; 9

Query Match 9.4%; Score 499; DB 2; Length 1134;

Best Local Similarity 23.5%; Pred. No. 1.2e-17;

Matches 193; Conservative 134; Mismatches 288; Indels 206; Gaps 29;

DNA Res. 5, 55-76, 1998
A;Title: Complete sequence and gene organization of the genome of a hyper-thermophilic a
A;Reference number: A71000; MUID:98344137; PMID:9679194
A;Accession: D71203
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-650 <KAW>
A;Cross-references: UNIPROT:O59524; GB:AP000007; NID:g3236134; PIDN:BAA31019.1; PID:g325
A;Experimental source: strain OT3
A;Note: this accession replaces an interim accession for a sequence replaced by GenBank
C;Genetics:
A;Gene: PH1896
C;Keywords: ATP; nucleotide binding; P-loop
F;33-40/Region: nucleotide-binding motif A (P-loop)
F;131-136/Region: nucleotide-binding motif B
F;135-138/Region: DEAD/H motif

Query Match 8.3%; Score 441.5; DB 2; Length 650;
Best Local Similarity 25.9%; Pred. No. 4.6e-15;
Matches 185; Conservative 122; Mismatches 261; Indels 147; Gaps 29;
QY 306 LQLRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKHDLDKDKKASEPGKVIVLV-N 364
Db 11 LQRLYQ-ELIYAKCKERNCLIVLPTGLGKTIAMMIADYRLDK-----YGGKVLMLAPT 64
QY 365 KVLVVE--QLFRKEFPFLKKWYRVICLSGDTQLKISFPPEVVKs---CDIIISTAQILEN 419
Db 65 KPLVLQHAETFRKFFTLPPKEK---IVALTGE---ISPNERVWARAKVVIATPQTIE 117
QY 420 SLLNLENGEDAGVQLSDFSLIIIDECHHTNKEAVYNNIMRHYLMQKLKNNRLKKNKPV 479
Db 118 DLL-----VGRISLEDVTLIIIDEAHRVAGNYAYVYIAKYLKQ-AKN----- 159
QY 480 PLPQILGLTASPGVGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQIQEPCKKF 539
Db 160 --PHVIGLTASPG-----STPEKIMEVLRNLGIEHYRSENSPDVKPYVQG--IKF 207
QY 540 AIADATREDPPEK--KLEIMTRIQTQYCMSPMSDFGTQPYEQWAIQMEKKAAGKGNR-- 595
Db 208 EWVKVELPELYKEVRKLLRDMRLR-----DSLKPLAEAGL--LDSTSADIPKKEILRAGQII 261
QY 596 -KERVCAEH-LRK-----YNEALQINDTIRMIDAY-----THLETFYNEEKDKKFAVIED 643
Db 262 NEEMAKGNHURKLLLFHAMALKLHAIELLETOGLSALRVYLRKLYEEAK----- 312
QY 644 DSDEGGDDYCDGDEDEDDLKPLKLDDETDRFLMTLFFENNKMMLKRLAENPEYENKLT 703
Db 313 -----AGSTRASKELFLDKRMKKALALLIQ-----AKELGIDHPKMEV 350
QY 704 LRNTIMEQYTRTEESARGIIFTKTROSAYALSQWITENKFAEVGVKAHHLIGAGHSSEF 763
Db 351 LKGLIREQLKR-KENSKIIVFTNYRETAKVW-----BELMKDGIKARRFVGQASREND 403
QY 764 KPMTQNEQKEVISKFRGTGKINLLIATTVAAEGLDIKECNIVIRYGLVTNEIAMVQARGR- 822
Db 404 RGMQSKEQKLILDARFARSEFNVLVATSVGEGLDVPEVDLVIFYEPVPSAIRSVQRRGT 463
QY 823 ARADESTYVLVAHSGSGVIEHETVNDFREKMYKAIHCQNMKPEEYAHKILELQMSIM 882
Db 464 GRQKPGRVILMAQGTDRDEAYYWSRQKEKIMRETIRMVSQM-----VRKERQLSLESYV 518
QY 883 EKKMKTKNIAKHYKNPNPSLITFLCKNCVSLACSGEDIHVIEKMHVNMTPEFKEL----- 938
Db 519 KREVNEE---IKBEKKETGGIKVVIDSRELRS-----EVVKKLTKLGKIEVRLDVGD 569
QY 939 YIVRENKALQKKCADYQING-----EIIKCGQAWGTMMVH 974
Db 570 YIVSEDVAIERKKSANDFIQSIIDGRFLDQVKRLKEAYPRPVIIIEGLYGIRNVH 624

RESULT 8
A69432
ATP-dependent RNA helicase homolog - Archaeoglobus fulgidus

C;Species: Archaeoglobus fulgidus
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: A69432
R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodso
.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A;Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S
Smith, H.O.; Woese, C.R.; Venter, J.C.
A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae
A;Reference number: A69250; MUID:98049343; PMID:9389475
A;Accession: A69432
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-741 <KLE>
A;Cross-references: UNIPROT:O28814; GB:AE001002; GB:AE000782; NID:g2689325; PIDN:AAB897
C;Superfamily: ATP-dependent RNA helicase eIF-4A
C;Keywords: ATP; nucleotide binding; P-loop
F;36-43/Region: nucleotide-binding motif A (P-loop)
F;132-137/Region: nucleotide-binding motif B
F;136-139/Region: DEAD/H motif

Query Match 7.9%; Score 417; DB 1; Length 741;
Best Local Similarity 23.8%; Pred. No. 9.9e-14;
Matches 169; Conservative 114; Mismatches 230; Indels 198; Gaps 30;
QY 303 EPELQLRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKHDLDKDKKASEPGKVIVL 362
Db 11 ENTIERRMYQISIAATALT-KNTLVVIPTGLGKTTIAALVIASRL-----LNEDGKVLFL 64
QY 363 VNKVLLVEQLFRKEFPFLKKWYRV---IGLSGDTQLKISFPPEVVKSCDIIISTAQILEN 419
Db 65 APTKPLVEQHAR-----FLKRVLKVEEIVSLSGEVPPE-KRKELWEKARIVVSTPQVVEN 118
QY 420 SLLNLENGEDAG-VQLSDFSLIIIDECHHTNKEAVYNNIMRHYLMQKLKNNRLKKNKPV 478
Db 119 DLL-----AGRISLEDVILVVFDEAHRVAGNYAYVYIAKEYLRTAKK----- 160
QY 479 IPLPQILGLTASPGVGATKQAKAEHILKLCANL--DAFTIKT----- 520
Db 161 ---PLILAMTASPG-----SDPERIMEVIQNLGIEAIEVTRTEWSDVAPYVGKKRIE 209
QY 521 -----VKENLDQLKNQIQEPCKKFAIADATR-----EDPFPKEKLLLEIMTRIQTQYCMSP 569
Db 210 WIKVDIPEEMKEVKERLKE-CIKIRFKRLRELWIEVENSCKRDLALQEAQAEASSQ 268
QY 570 MSDFGTQPYEQWAIQMEKKAAGKGNRKERVCAEHLRKVNEALQINDT--IRMDAYTHLE 627
Db 269 SSEI-----FEALSILAE-----IMKLQHAVELIETQGVKAVKSYLK-- 305
QY 628 TFYNEEKDKKFAVIEDSDSDEGGDEYCDGDEDEDDLKPLKLDDETDRFLMTLFFENNKMML 687
Db 306 -----KLVREATSKGG-----SKAA 320
QY 688 KRLAENP-----EYNEKLTKLNRNTIMEQYTRTEESARGIIFTKTROSAYALS 735
Db 321 KSIIVGDPFIFKKAVIALSKCKVEHPKLEKLEILKEQEKNPDS-RVIVFTNYRDSAEML- 378
QY 736 QWITENEKFAEVGVKAHHLIGAGHSSEFKPMTQNEQKEVISKFRGTGKINLLIATTVAAEEG 795
Db 379 --VNELSPLFPVA---KFVQGASRDNDKGMQRQKEQIETIDKFRRGVYKVLVATSVGEEG 432
QY 796 LDIKECNIVIRYGLVTNEIAMVQARGR-ARADESTYVLVAHSGSGVIEHETVNDFREKMM 854
Db 433 LDIPSTDVLVVFYEAVPSEIRAIQRKGRGTGRGREGRIVVLVTKGT-----RDEAY 481
QY 855 YKAIHCQNMKPEEYAH-KILELQMSIMEKKMKTGRNIAKHYKNPNPSLITFLCKNCVSL 913
Db 482 Y-----YSSMKKERKMYDKILEI--KRIIDRK---QRSIGDYVLPEETGIKIVDSREL 531
QY 914 ACSGEDIHVIEKMHVNMTPEFKEL-----YIVRENKALQKKCADYQINGEI 960
Db 532 S-----EVVKHLREIGAKIERNLEVADYVVSVDRAVERKTVEDFLNSII 576

QY 574 GTQPYEQWAIQMEKKAAGKGNRKERVCAEHLRKYNEALQINDTIRMIDA-----YTHLE 627
Db 269 -----DEEVYELIKVCSEALKMHAKELLESQKSVFLNYIN 306
QY 628 TFYNEEKDKKFAVIEDDSDEGGDDEYCDGDEDDLKPLKLDDETDRFLMTLFFENNMKL 687
Db 307 KLSMQRTKSAKSIYNDE-----KVBRAVNLLMKSDVEHPKL- 342
QY 688 KRLAENPEYENEKLTCLRNTIMEQYTRTEESARGIIFTKTRQSAAYALSQWITENEKEFAEV 747
Db 343 -----GKVVDVMVKNILEK---NKDERIIIFAQYRDTVEKIVNLLTON----- 381
QY 748 GVKAAHHLIGAGHSSEFKPMTQNEQKEVISKFRGTGKINLLIATTVAEEGLDIKECNIVIRY 807
Db 382 GIKAIRFIGQA-NKECKGMSQKEIQEIAIERFKK-EGSVLVSTSVSEEGIDIPSVNYIIFY 439
QY 808 GLVTNEIAMVQARGARADE--STYVLVAHSGSGVIEHETVNDPFREKMMYKAIHCVQNM- 864
Db 440 EPVPSEIRFIQRRGRAMRGEGGVYVLIAGTADAAYRSAL-YKEREMKRL----KNMC 495
QY 865 -----KPEEYAHKILELQMSIMEKMKTKRNIAXHYK-----NNPSLITFLCKN 909
Db 496 YLLNKRLOKQFEKSKSEIEKEETEEIKEKIESKTAVKEETKEHEEKTCKPVTILDFTIKQ 555
QY 910 CSV--LACSGED-----IHVIEKMHVNMT------EFKEL-----YIVRE 943
Db 556 IEVKERSKSEEDKIQEIKPKPIKIIVDVREKNMAKLLHNYANIELKTLEVGDYVLSD 615
QY 944 NKALQKKCADYQINGEI 960
Db 616 RVVVERKTAEDFVNSII 632

RESULT 11

S44849
K12H4.8 protein - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 14-Sep-1994 #sequence_revision 12-May-1995 #text_change 30-Jun-2001
C;Accession: S44849
R;Favello, A.D.
submitted to the EMBL Data Library, May 1993
A;Description: Sequence of the C. elegans cosmid K12H4.
A;Reference number: S44851
A;Accession: S44849
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1822 <FAV>
A;Cross-references: EMBL:L14331; NID:g289702; PIDN:AAA28101.1; PID:g289703
C;Genetics:
A;Introns: 17/3; 72/1; 116/3; 170/1; 216/3; 340/3; 401/3; 444/3; 512/2; 634/3; 694/3; 74
C;Keywords: ATP; nucleotide binding; P-loop
F;33-40/Region: nucleotide-binding motif A (P-loop)
F;141-146/Region: nucleotide-binding motif B
F;145-148/Region: DEXH motif

Query Match 6.6%; Score 350; DB 2; Length 1822;
Best Local Similarity 23.5%; Pred. No. 9.1e-10;
Matches 175; Conservative 118; Mismatches 249; Indels 202; Gaps 37;

QY 309 RPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKD------LDKKKASEPGKVIV 361
Db 14 RDYQVELLDKATK-KNTIVQLGTGSGKTFIAVLLKEYGVQLFAPLDQGGK-----RAFF 67
QY 362 LVNKKVLLVEQLFRKEFPFLKKWYRVIGLSGDT-----QLKISFPEVVKSCDIIISTAQI 416
Db 68 VVEKVNLV EQ--QAIHIEVHTSFKVGVGHGQTSSGLWDSKEQCDQFMKRHHVVVITAQC 124
QY 417 LENSLLNLENGEDAGVQLSDFSLLIIDECHHT-NKEAVYNNIMRHYLMQKLKNRLKKN 475
Db 125 L-----LDLIRHAYLKIEDMCVLIFFECHHALGSGHPYRSIMVDY-----KLLKKD 170
QY 476 KPVIPLPQILGLTASPGVGGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQIQEP 535

Db 171 KPV---PRVLGLTASL-IKAKVAPEKLMQKLESAMDS-VIETASD-LVLSLSKYGAKP 224
QY 536 -----CKKFAIA--DATREDPFKEKLLLEIMTRIQTICYQMSPMDSFGTQ----- 576
Db 225 YEVVIIICKDFEIGCLGIPNFDTVIEIFDETVAFVNTTTTEHPDLDLDPRRPIKDSLKTTR 284
QY 577 -----PYEQWAIQMEKKAAGKGNRKERVCAEHLRKYNEALQIND-TIRMID-AYTHLE 627
Db 285 AVFRQLGPWAAW-----RTAQWEEKELGKIISQVLPDKTLRFLNMAKTSMI 331
QY 628 TFYNEEKDKKFAVIEDDSDEGGDDEYCDGDEDDLKPLKLDDETDRFLMTLFFENNMKL 687
Db 332 TI-----KRL------EPEMKKIKSIEALRPYVPQVRIRLFEIL 365
QY 688 KRLAENPEYENEKLTCLRNTIMEQYTRTEESARGIIFTKTRQSAAYAL-----SQWITE 740
Db 366 ETF--NPEFQKERM-KLEA-----EHLSAIIFVDQRYAIAYSLLLMMRHIKSW--- 410
QY 741 NEKFAEVGVKAHHLIGAG---HSSEFKPMTQNEQKEVISKFRGTGKINLLIATTVAEEGL 796
Db 411 EPKFK--FVNPDYVVGASGRNLASSDSQGL-HKRQTEVLRFRHRNEINCLIAATSVLEEGV 467
QY 797 DIKECNIVIRYGLVTNEIAMVQARGARADESTYVLVAHSG-----SGVIE 842
Db 468 DVKQCNLVIKFDRPLDMRSYVQSGKRARRAGSRVYITVEEKDTAAAYCSKLPSDIFTRLVP 527
QY 843 HETVNDPFREKMMYKAIHCVQNMKP--EYAHKILELQMSIMEKMKTKRNIAXHYKNP 900
Db 528 HNQIPIEENGVT- -YCAELLLPINSPIKHA- - - - -VLKNPMPNPKT- - - - - 569
QY 901 SLITFLCKNCVSLACSGE-DIHVIEK- - - - -MHVNMTPE- - - - -FKE 937
Db 570 AQMAVALEACRQLHLEGELDDNLLPKGRESIAKLEHIDEEDPEYAPGIAAKVGSSKRKQ 629
QY 938 LYIVRENKAL- - - - -QKKCADY 954
Db 630 LYDKKIARALNESFVEADKECFIY 653

RESULT 12

E69055
ATP-dependent RNA helicase, eIF-4A family - Methanobacterium thermoautotrophicum (strai
C;Species: Methanobacterium thermoautotrophicum
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: E69055
R;Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiواني, N
ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997

A;Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: func
A;Reference number: A69000; MUID:98037514; PMID:9371463

A;Accession: E69055

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-738 <MTH>

A;Cross-references: UNIPROT:O27466; GB:AE000903; GB:AE000666; NID:g2622514; PIDN:AAB858

A;Experimental source: strain Delta H

C;Genetics:

A;Gene: MTH1415

C;Superfamily: ATP-dependent RNA helicase eIF-4A

C;Keywords: ATP; nucleotide binding; P-loop

F;38-45/Region: nucleotide-binding motif A (P-loop)

F;134-139/Region: nucleotide-binding motif B

F;138-141/Region: DEXH motif

Query Match 6.4%; Score 342; DB 1; Length 738;
Best Local Similarity 24.7%; Pred. No. 6.5e-10;
Matches 172; Conservative 106; Mismatches 234; Indels 184; Gaps 31;

QY 302 PEPELQLRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKDHLDKKASEPGKVIV 361
Db 12 PE-KIEARTYQQLLAADVIRKGNMIVAPTALGKTVVAVLVAERLRKYRGS----KVLI 66

QY 362 L-VNKVLVEQLFRKEFPFLLKKWYRVICLSGDTQLKISF-PEWK-----SCDIIISTAQ 415
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 67 LSPSKPLAIQH--ESSFREPM-----LATCTSLTSGIKPSEKERWERIKSQIISATPQ 116
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

QY 416 ILENSLLNLEDAG-QVLSDFSFLIIDCHHTNKEAVNNIMRHYLMQMKLKNRLKKE 474
: : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 117 TVESDI-----AGRYDLRDVSLIVDFECHRAGVSYSVFILASNY----IQAR--- 161
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

QY 475 NKPVIPLPOILGLTASPVGGA TKQA KAEHILKLCANL--DAFTIKT-----VKENLDQ 527
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 162 -----HPLILGLTASPGAD-----EDKI TVCENLFMNEVVVKTEGD PDPVPYLXP 207
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

QY 528 LKNQIQEPCKFAIADATREDPFKEKLEIMTRIQTQCMSMSDFGTQPYEQWA IQMEK 587
: : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 208 IKIEWVK-----VRMTPELEDIRELLRKVLKN-RLKMLKNLGVID----TISVGK 252
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

QY 588 KAAKKG-----NRKERVCAEHL RKYNEALQINDTIMIDAYTHLET FYNEEKDKKFAVIE 642
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 253 KDILLKARGRVQNRIARTSTPPRACYRAISLLASCINVEHALELET----- 298
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

QY 643 DDSDEGGDDEYCDGEDDDLKKPLKLDTRDFLMTLFFENNKM LKR LAENPEY----- 696
: : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 299 ----QG-----IRPL-----HQY LLRLKEKT KA AKGLLADPD FTRAMHL 334
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

QY 697 -----ENEKLTCLRNTIMEQYTRTEESARGIIFTKRSAYALSQWITENEKPAEV 747
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 335 TRRAMMSGVHPKLDRLME-I LR ELKGDE-A RI IVTFQRDTLEE I Y Q-----RCKRE 386
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

QY 748 GVKAHH ILIGAGHSSEFKPM TONEQE VISKFRTGK INLLIATTVAEEGLDIKECNIVIRY 807
: : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 387 GINAVFYGONS RS GEK GLTKQQORD II KS FRMGNDHVLLSTS VA EEG IDIPSV D LVVMY 446
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

QY 808 GLVTNEIAM VQARGRARADESTYLVLA HSGS GV IEHETV ND FRE KM MYK AIHC VQNM KP E 867
: : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 447 EPVPSEIRMIQRRGR TG RKRG RMV-----LITEK TRD---EAYYYSS IRKERSMK-E 496
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

QY 868 EYAHKITELQMQSIM EK MKTKRNIAKH YKNNPSLI TF LC KNCS VLACSGEDI HVIEKMH 927
: : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 497 NL RGGSVNVEVNP IMEP-----SGEG PF IY ADSR 525
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

QY 928 HVN--MTPEFKELYIVRENKALKQKCADYQINGEI 961
|| : || : || : || : || : || : || : || : || : || : || : || : || : || : || : || :
Db 526 EVNSRVLR ELKKIGVD FELK PL--AVGDYQISEDTI 559
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

RESULT 13
D84386
ATP-dependent RNA helicase homolog eIF-4A [imported] - Halobacterium sp. NRC-1
C;Species: Halobacterium sp. NRC-1
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: D84386
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Lasky,
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Jablo
A;Title: Genome sequence of Halobacterium species NRC-1.
A;Reference number: A84160; MUID:20504483; PMID:11016950
A;Accession: D84386
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-784 <STO>
A;Cross-references: UNIPROT:Q9HMW5; GB:AEO04437; NID:g10581766; PIDN:AAG20456.1; GSPDB:
C;Genetics:
A;Gene: eif4a
C;Superfamily: ATP-dependent RNA helicase eIF-4A

Query Match 5.9%; Score 313.5; DB 2; Length 784;
Best Local Similarity 21.6%; Pred. No. 2e-08;
Matches 126; Conservative 91; Mismatches 190; Indels 175; Gaps 22;

QY 306 QLRLPYQMEVAOPALEGNIIIICLP TSGSKTRVAVYAKDHLDKKKASEPGKVIVLNK 365

```

18 IEARQYQLQIAAAARQG-HTLVCLPTGLGKTTVSLLVTAAYRL-----ADDAGGTALLLAPT 72
366 VLLVEQLFRKEFPFLKKWYR-----VIGLSGDTQLKISFPEVVKs-----CDIIIS 412
73 KPLVEQ-----HAGFYREALAIPDDDDVVVFTGETR-----PDRRAAWTDARVVVA 118
413 TAQILENSLLNLENGEDAGVQLSDFSLLIIIDCHHTNKAEAVYNNIMRHYLMQKLKNRLK 472
119 TPQVVENDLVG-----GRIDMDDVVHCTFDECHRATGDYAYTYVAERYHAD----- 164
473 KENKVPVILPQILGLTASPGVGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQI 532
165 -----AAPLVTAMSASPG-----GTEAEIRTVCENTGVGNVEVMTED----- 202
533 QEPCKKFAIADATREDPFKEKLLLEIMTRIQTICYQMSPMDSFGTQPYE---QWAIQMEKKA 589
203 -----DA-----DVGEHTHDTDVQW----- 217
590 AKKGNRKERVCAEHLRKYNEALQINDTIRMI--DAYTHLETF-----YNEE 633
218 -----ERVTL-----EEILEVRDAINDVIEDRLAKLREIGVTKASSPDISQKDLNEI 265
634 KDKKFAVIEDSDSDEGGDDEYCDGDEDEDLKKPLKLDETDRF-LMTLFFEN----- 683
266 RARLQQLIDDDSDGYQGMsvHAeVMK--LKRAVELVETQSVESVRRYFERQORNAANTAG 323
684 -NKMLKRLAENPEYENEKLT-----KLKNT--IMEQYTRTEESARGIIFTKTRQSA 731
324 ASKASQRLVSEPAVKRAMRTAREFDGLHPKFRQARMLLAETLGIEDGDRVIVFTESRDTA 383
732 YALSQWITENEKFAEVGVKAHHLLIGAGHSSEFPKPMTQNEQKEVISKFRGTGINLLIATTV 791
384 EALTAFLGEH-----FDRRRFVGGQDADGSDGMTQTEQRETLLAEFRNGDFEVLVSTSV 436
792 AEEGLDIKECNIVIRYGLVTNEIAMVQARGARADESTYVLV 833
437 AEEGLDVPEVDLVLFEPVPPTAIRSVQRKGTGRQTAGRvvv 478

```

RESULT 14

S48436
 probable RNA helicase YIR002c - yeast (*Saccharomyces cerevisiae*)
 N;Alternate names: hypothetical protein YIB2c
 C;Species: *Saccharomyces cerevisiae*
 C;Date: 02-Dec-1994 #sequence_revision 02-Dec-1994 #text_change 09-Jul-2004
 C;Accession: S48436; S50885
 R;Badcock, K.; Churcher, C.
 submitted to the EMBL Data Library, August 1994
 A;Reference number: S48432
 A;Accession: S48436
 A;Molecule type: DNA
 A;Residues: 1-993 <BAD>
 A;Cross-references: UNIPROT:P40562; GB:Z47047; EMBL:Z38062; NID:G603997; PID:G763347; M
 R;Voss, H.; Tamames, J.; Teodoru, C.; Valencia, A.; Sensen, C.; Wiemann, S.; Schwager, M
 yeast 11, 61-78, 1995
 A;Title: Nucleotide sequence and analysis of the centromeric region of yeast chromosome
 A;Reference number: S50795; MUID:95282515; PMID:7762303
 A;Accession: S50885
 A;Status: nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-993 <VOS>
 A;Cross-references: EMBL:X79743
 A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1994
 C;Genetics:
 A;Gene: SGD:MPH1
 A;Cross-references: SGD:S0001441
 A;Map position: 9R
 C;Keywords: ATP; nucleotide binding; P-loop
 F;107-114/Region: nucleotide-binding motif A (P-loop)
 F;205-210/Region: nucleotide-binding motif B
 F;209-212/Region: DEAD/H motif

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 27, 2005, 21:02:13 ; Search time 6327 Seconds
(without alignments)
5903.386 Million cell updates/sec

Title: US-09-515-363C-2
Perfect score: 5311
Sequence: 1 MSGYSTDENFRYLISCFRA.....LPITFPNLDYSECCLFSDED 1025

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09515363/runat_26012005_152043_6148/app_query.fasta_1.1223
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09515363 @CGN_1_1_5180 @runat_26012005_152043_6148 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*

1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2673	50.3	2304	3 AK037057	AK037057 Mus muscu
2	2633	49.6	2331	3 AK040519	AK040519 Mus muscu
3	1597	30.1	1013	5 BQ233683	BQ233683 AGENCOURT
4	1376	25.9	870	5 BQ960157	BQ960157 AGENCOURT
5	1289.5	24.3	918	5 BU189982	BU189982 AGENCOURT
6	1257	23.7	1174	5 BU902097	BU902097 AGENCOURT
7	1218	22.9	1115	4 BM467983	BM467983 AGENCOURT
8	1206.5	22.7	1035	2 BF337464	BF337464 602035195
9	1129	21.3	729	4 BF983236	BF983236 602305873

10	1122	21.1	972	6	BY720783	BY720783
11	1122	21.1	1197	3	AK018602	AK018602 Mus muscu
12	1098	20.7	827	4	BG741146	BG741146 602631817
13	1094.5	20.6	781	2	BF686405	BF686405 602143786
14	1081.5	20.4	1041	2	BE882040	BE882040 601505326
15	1078	20.3	627	5	BX492926	BX492926 DKFZp781C
16	1066	20.1	1046	4	BM476961	BM476961 AGENCOURT
17	1059	19.9	755	5	BQ772836	BQ772836 UI-H-FE0-
18	1055.5	19.9	1239	4	BM467774	BM467774 AGENCOURT
19	1045	19.7	4098	3	BC070029	BC070029 Homo sapi
20	1021.5	19.2	800	7	CO396840	CO396840 AGENCOURT
21	1016	19.1	672	5	BQ316075	BQ316075 CM3-CT027
22	1016	19.1	672	5	BQ316108	BQ316108 CM3-CT027
23	1016	19.1	672	5	BQ316120	BQ316120 CM3-CT027
24	994	18.7	671	6	CB453859	CB453859 709752 MA
25	990.5	18.6	911	4	BI454996	BI454996 603173580
26	990	18.6	672	6	CA339255	CA339255 NISC_lx12
27	989.5	18.6	1611	4	BM455238	BM455238 AGENCOURT
28	954	18.0	682	7	CF363618	CF363618 833509 MA
29	952.5	17.9	932	4	BG967530	BG967530 602833306
30	947	17.8	647	6	CB424267	CB424267 598511 MA
31	930.5	17.5	698	7	CO748417	CO748417 SNESTbaa6
32	928.5	17.5	701	6	CA423868	CA423868 UI-H-FE1-
33	911.5	17.2	568	5	BQ311714	BQ311714 QV3-BN004
34	900	16.9	2398	3	BC027369	BC027369 Mus muscu
35	899.5	16.9	784	6	CB237812	CB237812 AGENCOURT
36	896.5	16.9	790	2	BF160649	BF160649 601768541
37	892	16.8	1201	4	BM551088	BM551088 AGENCOURT
38	891.5	16.8	617	6	CB452989	CB452989 707909 MA
39	857.5	16.1	666	2	BB148072	BB148072 BB148072
40	855	16.1	619	5	BQ016388	BQ016388 UI-H-DT1-
41	841.5	15.8	787	4	BI453288	BI453288 603173001
42	838.5	15.8	573	6	CB443396	CB443396 694276 MA
43	836.5	15.8	596	2	BF146945	BF146945 uy26c05.x
44	834.5	15.7	680	7	CF365022	CF365022 835621 MA
45	830.5	15.6	761	2	BB627907	BB627907 BB627907

ALIGNMENTS

RESULT 1

AK037057

LOCUS

DEFINITION

Mus musculus adult female vagina cDNA, RIKEN full-length enriched library, clone:9930105B04 product:similar to MELANOMA DIFFERENTIATION ASSOCIATED PROTEIN-5 [Homo sapiens], full insert sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

AK037057 2304 bp mRNA linear HTC 03-APR-2004

Mus musculus adult female vagina cDNA, RIKEN full-length enriched library, clone:9930105B04 product:similar to MELANOMA DIFFERENTIATION ASSOCIATED PROTEIN-5 [Homo sapiens], full insert sequence.

AK037057 GI:26331913

HTC; CAP trapper.

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

1

Carninci, P. and Hayashizaki, Y.

High-efficiency full-length cDNA cloning

Meth. Enzymol. 303, 19-44 (1999)

99279253

10349636

2

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,

Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new genes

Genome Res. 10 (10), 1617-1630 (2000)

20499374

11042159

3

Shibata, K., Itoh, M., Aizawa, K.,

Nagaoka, S., Sasaki, N., Carninci, P.,

Konno, H., Akiyama, J., Nishi, K.,

Kitsunai, T., Tashiro, H., Itoh, M.,

Sumi, N., Ishii, Y., Nakamura, S.,

Hazama, M., Nishine, T., Harada, A.,

Yamamoto, R., Matsumoto, H.,

Sakaguchi, S., Ikegami, T.,

Kashiwagi, K.,

Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861
4
The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5
The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 2304)
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saitoh,K., Shibata,K., Shinagawa,A., Shiraki,T., Sano,H., Sasaki,D., Shibata,K., Sakai,C., Sakai,K., Sakazume,N., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.jp/
URL:http://fantom.gsc.riken.jp/.
Location/Qualifiers
1. .2304
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:9930105B04"
/db_xref="taxon:10090"
/clone="9930105B04"
/sex="female"
/tissue_type="vagina"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
224.-.2302
/note="unnamed protein product; putative similar to MELANOMA DIFFERENTIATION ASSOCIATED PROTEIN-5 [Homo sapiens] (SPTR|Q9H3G6, evidence: FASTY, 75.6%ID, 67.6%length, match=2079)"
/codon_start=1
/protein_id="BAC29687.1"
/db_xref="GI:26331914"
/translation="MSIVCSAEDSFRNLILFFRPRLKMVIQVEPVLHLIFLSAETKE QILKINTCGNTSAAELLSTLEQGWPLGTMQFVBALEHSGNPLAARYVKPTLTDL PSPSSETAHDECLHLLTLTLOPTLVLDKLLINDVLTDFEKGLLTVEDRNRISAAGNSGN ESGVRELLRRIVQENWFSLDVLRLQTNLDALFQELTGGCPEDNTDLANSSSHRRDGP AANECLLPAVDESSLETEAWNVDLILPEASCTDSSVTTESDTSIAEGSVSCFDESIGH

NSNMGRDSTMGSDSDSESVIOTKRVSPPELOLRPRPYOMEVAQPALDGKNIICLPTGSGKTRVAVYITKDHLDDKKQASGSKVIVLVNKMVLABOLFRKEFPNYPYKWKYRIGLSGDTQLKISPEPVVKSVDYVVIISTAIQILENSLLNLESGDDGVQLSDFSLIIDECHHTNKEAVYNNIMRRYLLKQKLRNNDLKQNKPAIPLQDGLTASPGVGAACKQSEAEKHILNICANLDAFTIKTVKENLQQLKHQIKEPKKFVIADDTRENPFKEKLLLEIMASIQTYCOKSPMSDFGTQHYEQWAIQMEKKAADGNRKDRVCAEHLRKYNEALQINDTIRMIDAYSHLETFFYDEKEKKFAVLNDSKXSDDEASSCNDQLKGDVYKSLKLDDETDFLMLNLFDDNKMKLKLAEEN"
ORIGIN
Alignment Scores:
Pred. No.: 2304
Score: 2673.00
Percent Similarity: 84.29%
Best Local Similarity: 75.65%
Query Match: 50.33%
DB: 3
US-09-515-363C-2 (1-1025) x AK037057 (1-2304)
QY 1 MetSerAsnGlyTyrSerThrAspGluAsnPheArgTyrLeuIleSerCysPheArgAla 20
Db 224 ATGTCGATTGTCGTCTTCAGAGGACAGTTCAGGAATCTCATCTTATCTTCAGGCC 283
QY 21 ArgValLysMetTyrIleGlnValGluProValLeuAspTyrLeuThrPheLeuProAla 40
Db 284 AGGCTGAAATGTACATTTCAGGTGGAGCCAGTGTGGACCCTCATCTTCTGTCTGCA 343
QY 41 GluValLysGluGlnIleGlnArgThrValAlaThrSerGlyAsnMetGlnAlaValGlu 60
Db 344 GAAACCAAGAGCAGATTCTTAAAGATCAACACCTGTGGTAACACAGCGCGGAGAA 403
QY 61 LeuLeuLeuSerThrLeuGluLysGlyValTrpHisLeuGlyTrpThrArgGluPheVal 80
Db 404 CTGCTGCTGAGCACCTTGGAGCAGGGACAATGGCCTCTGGGATGGACGAGATGTTCTG 463
QY 81 GluAlaLeuArgArgThrGlySerProLeuAlaAlaArgTyrMetAsnProGluLeuThr 100
Db 464 GAGGCCCTAGAGCACAGTGGCAATCCCTTAGCGCGCGCTATGTCAAACCCACACTCACT 523
QY 101 AspLeuProSerProSerPheGluAsnAlaHisAspGluTyrLeuGlnLeuLeuAsnLeu 120
Db 524 GATCTGCCCTCTCCTTCCTCTGAGACTGCCCATGACGAGTGTCTCCACTGTCTGACCCTC 583
QY 121 LeuGlnProThrLeuValAspLysLeuLeuValArgAspValLeuAspLysCysMetGlu 140
Db 584 CTCAGCCCCACTTTGGTGGACAAACTTCTGATTAACGATGTCTTGGACACTTGTTCGAG 643
QY 141 GluGluLeuLeuThrIleGluAspArgAsnArgIleAlaAlaGluAsnAsnGlyAsn 160
Db 644 AAGGACTATTGACAGCTCGAAGACAGAAATCGGATTTCTGCTGCAGGAACACAGCGGAAT 703
QY 161 GluSerGlyValArgGluLeuLysArgIleValGlnLysGluAsnTrpPheSerAla 180
Db 704 GAGTCAGGTGTAAGAGAGCTGCTGAGAGGATGTGCAGAAAGGAAACTGGTTTTCTTCTACC 763
QY 181 PheLeuAsnValLeuArgGlnThrGlyAsnAsnGluLeuValGlnGluLeuThrGlySer 200
Db 764 TTCTCGATGTTCTGCGCCAAACTGGAAATGATGACTATTCCAGAACTAAACAGGTGGA 823
QY 201 AspCysSerGluSerAsnAlaGluIleGluAsnLeuSerGlnValAspGlyProGlnVal 220
Db 824 GGCTGCCCAGAGACACACAGACTTGGCTAACTCGTCTCACAGAGATGGGCTGCAGCT 883
QY 221 GluGluGlnLeuLeuSerThrThrValGlnProAsnLeuGluLysGluValTrpGlyMet 240
Db 884 AATGAGTGTTCTTCTGCCTGTCGATGAGTCAACTCAACTTCCAGAGATGGGCTGGAACGTA 943
QY 241 GluAsnAsnSerSerGluSerSerPheAlaAspSerSerValValSerGluSerAspThr 260
Db 944 GACGACATATTACCAGAGGGCTTCTTGTACAGATTGCTGTGTGACCACAGAAATCAGACACA 1003
QY 261 SerLeuAlaGluGlySerValSerCysLeuAspGluSerLeuGlyHisAsnSerAsnMet 280

Db 1004 AGTTGGCAGAAAGAAAGTGTGAGTCTGCTTCGATGAAAGTCTTGGACATAACAGCAACATG 1063

QY 281 GlySerAspSerGlyThrMetGlySerAspSerAspGluGluAsnVal---AlaAlaArg 299

Db 1064 GGCAGGATTTCAGGCACCATGGGAAGTGATTCAGATGAAAGTGTCATCCAGACAAAAGA 1123

QY 300 AlaSerProGluProGluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAla 319

Db 1124 GTATCCCCGAGCCAGAACTGCAGCTCAGGCCTTACCAATGGAAGTGGCCCAACAGCT 1183

QY 320 LeuGluGlyLysAsnIleIleCysLeuProThrGlySerGlyLysThrArgValAla 339

Db 1184 CTAGATGGGAAGAAATATTATTATCTGCCTCCCCACGGGAGTGGGAAAACCCAGAGTGGCT 1243

QY 340 ValTyrIleAlaLysAspHisLeuAspLysLysLysLysAlaSerGluProGlyLysVal 359

Db 1244 GTTTACATCACCAAGATCACTTAGACAAGAAGAAGCAGGCATCTGAATCCGGGAAGGTT 1303

QY 360 IleValLeuValAsnLysValLeuLeuValGluGlnLeuPheArgLysGluPheGlnPro 379

Db 1304 ATCGTTCTTGTCAATAAGGTAATGTTAGCAGAACAACTTTTCCGAAAAGAGTTCAACCCA 1363

QY 380 PheLeuLysLysTyrTyrArgValIleGlyLeuSerGlyAspThrGlnLeuLysIleSer 399

Db 1364 TATTGGAAGAAATGGTATCGAATTATTGGATTAAAGTGGCGATACCAGCTGAAAATATCA 1423

QY 400 PheProGluValValLysSerCysAspIleIleIleSerThrAlaGlnIleLeuGluAsn 419

Db 1424 TTTCCAGAAGTTGTCAAATCTTACGATGTTATTATCAGCACTGCTCAAAATCCTTGAAAAC 1483

QY 420 SerLeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPheSerLeu 439

Db 1484 TCCCTCTTAAATCTGGAGAGTGGAGACGATGACGGTGTGCAGCTGCAGACTTCTCTCTC 1543

QY 440 IleIleIleAspGluCysHisHisThrAsnLysGluAlaValTyrAsnAsnIleMetArg 459

Db 1544 ATTATCATTTGATGAGTGCCATCACACCAAGAAGGAGGCAGTCTATACAACATCATGAGA 1603

QY 460 HisTyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysLysGluAsnLysProValIle 479

Db 1604 CGATATTTGAAGCAGAAAGCTGAGAAACAATGACCTCAAGAAACAACAAACCCAGCCATT 1663

QY 480 ProLeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyGlyAlaThrLysGln 499

Db 1664 CCCCTGCCGAGATACTAGGACTGACAGCTTCACCTGGTGTGGAGCAGCCAAAAGCAG 1723

QY 500 AlaLysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThrIleLys 519

Db 1724 TCTGAGGCTGAAAACATATTTTAAATATATGTGCCAAATCTTGATGCCTTTACCATTTAA 1783

QY 520 ThrValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluProCysLysLysPhe 539

Db 1784 ACAGTGAAGAGAAATCTTGGTCAACTCAACACCAATAAAGGACCATGCAAGAAATTT 1843

QY 540 AlaIleAlaAspAlaThrArgGluAspProPheLysGluLysLeuGluIleMetThr 559

Db 1844 GTGATTGCTGATGACACAGAGAAAATCCATTTAAAGAGAAACTTCTAGAAATTTATGGCA 1903

QY 560 ArgIleGlnThrTyrCysGlnMetSerProMetSerAspPheGlyThrGlnProTyrGlu 579

Db 1904 AGCATTGAGACTTACTGCCAAAAAAGTCCAAATGTCTGAGATTTTGGAAACCCCAACATTATGAG 1963

QY 580 GlnTrpAlaIleGlnMetGluLysLysAlaAlaLysLysGlyAsnArgLysGluArgVal 599

Db 1964 CAGTGGGCCATTCAATGGAGAAAAAAGCTGCTAAAGACGGAATTCGCAAGATCGCGTC 2023

QY 600 CysAlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMet 619

Db 2024 TGTGCAGAACATTTTGAGGAAGTACAACGAAGCCCTACAAATCAACGACACGATCCGAATG 2083

QY 620 IleAspAlaTyrThrHisLeuGluThrPheTyrAsnGluGluLysAspLysLysPheAla 639

Db 2084 ATTGATGCATATAGCCACCTGGAGACATTTCTACACTGATGAGAAAGAAAGTTTCGCA 2143

QY 640 ValIleGluAspAspSerAspGluGlyGlyAspAspGluTyrCysAspGlyAspGluAsp 659

Db 2144 GTCCTC---AATGACAGCGACAAGAGTGTGATGACGAGGCCAGCACTGCAATGACCAACTT 2200

QY 660 GluAspAspLeuLysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeu 679

Db 2201 AAGGGCGATGTAAGAAATCTTTTGAAACTGGACGAAACGGATGAATTTCTCATGAATTG 2260

QY 680 PhePheGluAsnAsnLysMetLeuLysArgLeuAlaGluAsn 693

Db 2261 TTCTTTGATAACAAGAAATGTTTGAANAACCTAGCTGAAAAC 2302

RESULT 2

AK040519

LOCUS

DEFINITION

AK040519 2331 bp mRNA linear HTC 03-APR-2004

Mus musculus 0 day neonate thymus cDNA, RIKEN full-length enriched library, clone:A430105A06 product:similar to MELANOMA DIFFERENTIATION ASSOCIATED PROTEIN-5 [Homo sapiens], full insert sequence.

ACCESSION

AK040519

VERSION

AK040519.1 GI:26333792

KEYWORDS

HTC; CAP trapper.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)

2 20499374

11042159

REFERENCE

3

SHIBATA,K., ITOH,M., AIZAWA,K., NAGAOKA,S., SASAKI,N., CARNINCI,P., KONNO,H., AKIYAMA,J., NISHI,K., KITSUNAI,T., TASHIRO,H., ITOH,M., SUMI,N., ISHII,Y., NAKAMURA,S., HAZAMA,M., NISHINE,T., HARADA,A., YAMAMOTO,R., MATSUMOTO,H., SAKAGUCHI,S., IKEGAMI,T., KASHIWAGI,K., FUJIWAKE,S., INOUE,K., TOGAWA,Y., IZAWA,M., OHARA,E., WATAHIKI,M., YONEDA,Y., ISHIKAWA,T., OZAWA,K., TANAKA,T., MATSUURA,S., KAWAI,J., OKAZAKI,Y., MURAMATSU,M., INOUE,Y., KIRA,A. and HAYASHIZAKI,Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000)

4 20530913

11076861

REFERENCE

5

THE RIKEN GENOME EXPLORATION RESEARCH GROUP PHASE II TEAM AND THE FANTOM CONSORTIUM. Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001)

6 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002)

7 6 (bases 1 to 2331)

ADACHI,J., AIZAWA,K., AKIMURA,T., ARAKAWA,T., BONO,H., CARNINCI,P., FUKUDA,S., FURUNO,M., HANAGAKI,T., HARA,A., HASHIZUME,W., HAYASHIDA,K., HAYATSU,N., HIRAMOTO,K., HIRAKA,T., HIROZANE,T., HORI,F., IMOTANI,K., ISHII,Y., ITOH,M., KAGAWA,I., KAGUKAWA,T., KATO,H., KAWAI,J., KOJIMA,Y., KONDO,S., KONNO,H., KOUDA,M., KOYA,S., KURIHARA,C., MATSUYAMA,T., MIYAZAKI,A., MURATA,M.,


```
QY 400 PheProGluValVallysSerCysAspIleIleSerThrAlaGlnIleLeuGluAsn 419
Db 1483 TTTCAGAAAGTTGTCAAATCTTACGATGTTATTATCAGCACTGCTCAAATCCTTGAAAAAC 1542

QY 420 SerLeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPheSerLeu 439
Db 1543 TCCCTCTTAAATCTGGAGAGTGGAGACGATGACGGTGTGCAGCTGTCTCTCTC 1602

QY 440 IleIleIleAspGluCysHisHisThrAsnLysGluAlaValTyrAsnAsnIleMetArg 459
Db 1603 ATTATCATTTGATGAGTGCCATCACACCAACAAGGAGGCAGTCTATAACAACATCATGAGA 1662

QY 460 HisTyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysLysGluAsnLysProValIle 479
Db 1663 CGATATTTGAAGCAGAAGCTGAGAAACAATGACCTCAAGAAACAACAACACCGCCATT 1722

QY 480 ProLeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyGlyAlaThrLysGln 499
Db 1723 CCCCTGCCGCAGATACTAGGACTGACAGCTTCACCTGGTGTGGAGCAGCCAAAAGCAG 1782

QY 500 AlaLysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThrIleLys 519
Db 1783 TCTGAGGCTGAAAAACATATTTTAAATATATGTGCCAATCTTGATGCTTTTACCATTAAA 1842

QY 520 ThrValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluProCysLysLysPhe 539
Db 1843 ACAGTGAAGAGAAATCTTGGTCAACTCAAAACACCAATAAAGGAACCATGCAAGAAATTT 1902

QY 540 AlaIleAlaAspAlaThrArgGluAspProPheLysGluLysLeuLeuGluIleMetThr 559
Db 1903 GTGATTGCTGATGACACCAAGAGAAATCCATTAAAGAGAAACTTCTAGAAATTTATGCA 1962

QY 560 ArgIleGlnThrTyrCysGlnMetSerProMetSerAspPheGlyThrGlnProTyrGlu 579
Db 1963 AGCATTGAGACTTACTGCCAAAAAAGTCCAAATGTCAGATTTTGGAAACCCCAACATTATGAG 2022

QY 580 GlnTrpAlaIleGlnMetGluLysLysAlaAlaLysLysGlyAsnArgLysGluArgVal 599
Db 2023 CAGTGGGCCATTCAAATGGAGAAAAAAGCTGTAAAGACGGAAATCGCAAGATCGCGTC 2082

QY 600 CysAlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAsnAspThrIleArgMet 619
Db 2083 TGTGCAGAACATTTGAGGAAGTACACGAAGCCCTACAAATCAACGACACGATCCGAATG 2142

QY 620 IleAspAlaTyrThrHisLeuGluThrPheTyrAsnGluGluLysAspLysLysPheAla 639
Db 2143 ATTGATGCATATAGCCACCTGGAGACATTTACACTGATGAGAAAGAAAGAAGTTCCGA 2202

QY 640 ValIleGluAspSerAspGluGlyGlyAspAspGluTyrCysAspGlyAspGluAsp 659
Db 2203 GTCCTC--AATGACAGCGACAAAGAGTGTACGAGGCCGACGAGTTCGAATGACCAACTT 2259

QY 660 GluAspAspLeuLysLysProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeu 679
Db 2260 AAGGGCGATGTAAAGAAATCTTTGAAACTGGACGAAACGGATGAATTTCTCATGAATTG 2319

QY 680 PhePheGluAsn 683
Db 2320 TTCTTTGATAAC 2331

RESULT 3
BQ233683
LOCUS
DEFINITION BQ233683 1013 bp mRNA linear EST 02-MAY-2002
AGENCOURT_7258428 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5786526
5', mRNA sequence.
ACCESSION BQ233683
VERSION BQ233683.1 GI:20415083
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1013)
```

```
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12875 row: g column: 07
High quality sequence stop: 721.
FEATURES
Location/Qualifiers
source
1..1013
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5786526"
/tissue_type="leiomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_71"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2.1 kb. "
ORIGIN
Alignment Scores:
Pred. No.: 3.1e-160 Length: 1013
Score: 1597.00 Matches: 322
Percent Similarity: 95.61% Conservative: 5
Best Local Similarity: 94.15% Mismatches: 9
Query Match: 30.07% Indels: 8
DB: 5 Gaps: 3
US-09-515-363C-2 (1-1025) x BQ233683 (1-1013)
QY 529 LysAsnGlnIleGlnGluProCysLysLysPheAlaIleAlaAspAlaThrArgGluAsp 548
Db 2 AAAAAACCAATACAGGAGCCATGCAAGAAGTTTGCCATTGCAGATGCAACCAGAGAAGAT 61
QY 549 ProPheLysGluLysLeuLeuGluIleMetThrArgIleGlnThrTyrCysGlnMetSer 568
Db 62 CCATTTAAAGAGAAACTTCTAGAAATAATGACAAGGATTCAAACTTATTGTCAAATGAGT 121
QY 569 ProMetSerAspPheGlyThrGlnProTyrGluGlnTrpAlaIleGlnMetGluLysLys 588
Db 122 CCAATGTCAGATTTTGGAACTCAACCTATGAACAATGGCCATTCAAATGGAAAAAAA 181
QY 589 AlaAlaLysLysGlyAsnArgLysGluArgValCysAlaGluHisLeuArgLysTyrAsn 608
Db 182 GCTGCAAAAGAGGAAATCGCAAGAACGTTGTTGTGCAGAACATTTGAGGAAGTACAAT 241
QY 609 GluAlaLeuGlnIleAsnAspThrIleArgMetIleAspAlaTyrThrHisLeuGluThr 628
Db 242 GAGGCCCTACAAATTAATGACACAATTCGAATGATAGATGCGTATCTCATCTTGAACCT 301
QY 629 PheTyrAsnGluGluLysAspLysLysPheAlaValIleGluAspAspSerAspGluGly 648
Db 302 TTCTATAATGAAGAGAAAGATAAGAAGTTTGAGTCATAGAAGATGATAGTAGAGGGT 361
QY 649 GlyAspAspGluTyrCysAspGlyAspGluAspGluAspLeuLysLysProLeuLys 668
Db 362 GGTGATGATGAGTATTGTTGATGGTGTGATGAAGATGAGGATGATTTAAAGAAACCTTTGAAA 421
QY 669 LeuAspGluThrAspArgPheLeuMetThrLeuPheGluAsnAsnLysMetLeuLys 688
Db 422 CTGGATGAACACAGATAGATTTCTCATGACTTTTATTTTGGAAAAACAATAAAATGTTGAAA 481
QY 689 ArgLeuAlaGluAsnProGluTyrGluAsnGluLysLeuThrLysLeuArgAsnThrIle 708
Db 482 AGGCTGGCTGAAAAACCCAGAAATATGAAAAATGAAAAGCTGACCAAAATTAAAGAAATACCATA 541
```


QY 709 MetGluGlnTyrThrArgThrGluGluSerAlaArgGlyLeilePheThrLysThrArg 728
|||||
Db 542 ATGGAGCAATATACTAGGACTGAGGAATCAGCACGAGGAATAATCTTTACAAAAACACGA 601

QY 729 GlnSerAlaTyrAlaLeuSerGlnTrpIleThrGluAsnGluLysPheAlaGluValGly 748
|||||
Db 602 CAGAGTGCAATGCGCTTTCCAGTGGATTACTGAAATGAAAAATTTGCTGAAATAGGA 661

QY 749 ValLysAlaHisHisLeuIleGlyAlaGlyHisSerSerGluPheLysProMetThrGln 768
|||||
Db 662 GTCAAAGCCCCACCATCTGATTGGAGCTGGACACAGCAGTGAGTTCAAACCCATGACACAG 721

QY 769 AsnGluGlnLysGluValIleSerLysPheArgThrGlyLysIleAsnLeuLeuIleAla 788
|||||
Db 722 AATGAACAAAAAGAAAGTCATTAGTAAATTTCCGACTGGAAAAATAAATCTGCTTATCGCT 781

QY 789 ThrThrValAlaGluGluGlyLeuAspIleLysGluCysAsnIleValIleArgTyrGly 808
|||||
Db 782 ACCACAGTGGCAGAAAGAAAGTCTGGATATTAAAGAAATGTAACATTGTTATCCGTTATGGT 841

QY 809 LeuValThrAsnGluIleAlaMetValGlnAlaArgGlyArgAlaArgAlaAspGluSer 828
|||||
Db 842 CTCGTAC-AATGAAATAGCCATGGTCCAGGCCCGTGGTCAGCCAGAGCTGATGAGAGC 900

QY 829 ThrTyrValLeuValAlaHisSerGlySerGlyValIleGluHisGluThr-----Val 846
|||||
Db 901 ACCTACGTCTCGTGTGCTCACGTGGGTCA-----GAGTATCGAACGTCGAGACATT 951

QY 847 AsnAspPheArgGluLysMetMetTyrLysAlaIleHisCysValGlnAsnMetLysPro 866
|||||
Db 952 AATGATTTCCGAGA-AAGATGATGTATAAG---CTATACATGGGGTTCAAAATGAAACCG 1007

QY 867 GluGlu 868
|||||
Db 1008 GAGGAG 1013

RESULT 4
BQ960157
LOCUS
DEFINITION BQ960157 870 bp mRNA linear EST 21-AUG-2002
AGENCOURT_8949486 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:6472264
5', mRNA sequence.
ACCESSION BQ960157
VERSION BQ960157.1 GI:22375635
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 870)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM14005 row: 0 column: 17
High quality sequence stop: 738.
Location/Qualifiers
1. .870
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6472264"
/tissue_type="leiomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_71"

/note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2.1 kb."

ORIGIN

Alignment Scores:
Pred. No.: 1.36e-136 Length: 870
Score: 1376.00 Matches: 279
Percent Similarity: 97.90% Conservative: 1
Best Local Similarity: 97.55% Mismatches: 2
Query Match: 25.91% Indels: 5
DB: 5 Gaps: 0

US-09-515-363C-2 (1-1025) x BQ960157 (1-870)

QY 592 LysGlyAsnArgLysGluArgValCysAlaGluHisLeuArgLysTyrAsnGluAlaLeu 611
:::|||||
Db 3 GAAGGAAATCGCAAGAAACGCTGTTTGTGCAGAAAT-TTGAGGAAGTACAATGAGGCCCTA 61

QY 612 GlnIleAsnAspThrIleArgMetIleAspAlaTyrThrHisLeuGluThrPheTyrAsn 631
|||||
Db 62 CAAATAATGACACAAATTCGAATGATAGATGCGTATACTCATCTTGAAACTTCTATAAT 121

QY 632 GluGluLysAspLysLysPheAlaValIleGluAspAspSerAspGluGlyGlyAspAsp 651
|||||
Db 122 GAAGAGAAAGATAAGAGTTTGCAGTCATAGAGATGATAGTAGAGGTGGTGTATGAT 181

QY 652 GluTyrCysAspGlyAspGluAspGluAspLysLysProLeuLysLeuAspGlu 671
|||||
Db 182 GAGTATTGTGATGGTGTAGAGATGAGGATGATTAAAGAAACCTTTGAAACTGGATGA 241

QY 672 ThrAspArgPheLeuMetThrLeuPhePheGluAsnAsnLysMetLeuLysArgLeuAla 691
|||||
Db 242 ACAGATAGATTTCTCATGACTTTATTTTTTGAACAAATAAATGTTGAAAAGCTGGCT 301

QY 692 GluAsnProGluTyrGluAsnGluLysLeuThrLysLeuArgAsnThrIleMetGluGln 711
|||||
Db 302 GAAACCAGCAATATGAAAATGAAAAGCTGACCAAAATTAAGAAATACCAATAATGGAGCAA 361

QY 712 TyrThrArgThrGluGluSerAlaArgGlyIleIlePheThrLysThrArgGlnSerAla 731
|||||
Db 362 TATACTAGGACTGAGGAATCAGCACGAGGAATAATCTTTACAAAAACACGACAGAGTGCA 421

QY 732 TyrAlaLeuSerGlnTrpIleThrGluAsnGluLysPheAlaGluValGlyValLysAla 751
|||||
Db 422 TATGCGCTTTCCAGTGGATTACTGAAAATGAAAAATTTGCTGAAGTAGAGTCAAAGCC 481

QY 752 HisHisLeuIleGlyAlaGlyHisSerSerGluPheLysProMetThrGlnAsnGluGln 771
|||||
Db 482 CACCATCTGATTGGAGCTGGACACAGCAGTGAGTTCAAACCCATGACACAGAATGAACAA 541

QY 772 LysGluValIleSerLysPheArgThrGlyLysIleAsnLeuLeuIleAlaThrThrVal 791
|||||
Db 542 AAAGAAGTCATTAGTAAATTTGCGACTGGAAAAATAAATCTGCTTATCGTACCACAGTG 601

QY 792 AlaGluGluGlyLeuAspIleLysGluCysAsnIleValIleArgTyrGlyLeuValThr 811
|||||
Db 602 GCAGAAGAAGGTCTGGATATTAAAGAATGTAACTGTTATCCGTTATGGTCTCGTCACC 661

QY 812 AsnGluIleAlaMetValGlnAlaArgGlyArgAlaArgAlaAspGluSerThrTyrVal 831
|||||
Db 662 AATGAAATAGCCATGGTCCAGGCCCGTGGTCGAGCCAGAGCTGATGAGAGCACCTACGTC 721

QY 832 LeuValAlaHisSer-GlySerGlyValIleGluHisGluThrValAsnAspPheArgG1 851
|||||
Db 722 CTGTTGCTCACAGTGGGTTCCAGGAGTTATCGAACGTCGAGACAGTTAATGATTTCCGAGA 781

QY 851 uLysMetMetTyrLysAlaIleHisCysValGlnAsnMet-LysProGluGluTyr--Al 870
|||||
Db 782 GAAGATGATGTATAAAGCTATACATTGTGTTCAAATAATGAAAACCCAGAGGAGTATGGGC 841

QY 870 aHisLysIle 873
|||||

Db 842 TCATAAGATT 851

RESULT 5

BU189982

LOCUS

DEFINITION BU189982 918 bp mRNA linear EST 04-SEP-2002

AGENCOURT_7950921 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:6173848

5', mRNA sequence.

ACCESSION BU189982

VERSION BU189982.1 GI:22703966

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 918)

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Plate: LLAM13546 row: a column: 17

High quality sequence stop: 556.

Location/Qualifiers

1. .918

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:6173848"

/tissue type="leiomyosarcoma"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH_MGC_71"

/note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.

Average insert size 2.1 kb. "

ORIGIN

Alignment Scores:

Pred. No.: 2,91e-127 Length: 918

Score: 1289.50 Matches: 269

Percent Similarity: 95.14% Conservative: 5

Best Local Similarity: 93.40% Mismatches: 6

Query Match: 24.28% Indels: 8

DB: 5 Gaps: 1

US-09-515-363C-2 (1-1025) x BU189982 (1-918)

QY 360 IleValLeuValAsnLysValLeuLeuValGluGlnLeuPheArgLysGluPheGlnPro 379

Db 2 ATAGTTCTTGTCATAAGGTACTGCTAGTTGAACAGCTCTTCCGCAAGGAGTTCCACCA 61

QY 380 PheLeuLysLysTrpTyrArgValIleGlyLeuSerGlyAspThrGlnLeuLysIleSer 399

Db 62 TTTTGAAGAAATGGTATCGTGTATTGGATTAAAGTGGTGATACCAACTGAAATATCA 121

QY 400 PheProGluValLysSerCysAspIleIleSerThrAlaGlnIleLeuGluAsn 419

Db 122 TTTCCAGAAGTTGTCAAGTCCTGTGATATTATTATCAGTACAGCTCAAATCCTTGAAAC 181

QY 420 SerLeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPheSerLeu 439

Db 182 TCCCTCTTAAACTTGGAAAATGGAGAAGATGCTGGTGTTCATTGTCTAGACTTTTCCCTC 241

QY 440 IleIleIleAspGluCysHisHisThrAsnLysGluAlaValTyrAsnAsnIleMetArg 459

Db 242 ATTATCATTTGATGATGTCATCATCACACCAACAAGAAGCAGTGATATAATAATCATGAGG 301

QY 460 HisTyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysLysGluAsnLysProValIle 479

Db 302 CATTATTTCATGCAGAGTTTGAAAAACAATAGACTCAAGAAAGAAAAACAACCAAGTATT 361

QY 480 ProLeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyGlyAlaThrLysGln 499

Db 362 CCCCTTCCTCAGATACTGGGACTAACAGCTTCACCTGGTGTGGAGGGGCCACGAAGCAA 421

QY 500 AlaLysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThrIleLys 519

Db 422 GCCAAAGCTGAAGAACAACATTTTAAAAACTATGTGCCAATCTTGATGCTATTACTATAAA 481

QY 520 ThrValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluProCysLysLysPhe 539

Db 482 ACTGTTAAAGAAAAACCTTGATCACTGAAAAACCAAAATACAGGACCATGCAAGAAGTTT 541

QY 540 AlaIleAlaAspAlaThrArgGluAspPropheLysGluLysLeuLeuGluIleMetThr 559

Db 542 GCCATTGCAGATGCAACCAGAGAAGATCCATTTAAAGAGAAACTTCTAGAAATAATGACA 601

QY 560 ArgIleGlnThrTyrCysGlnMetSerProMetSerAspPheGlyThrGlnProTyrGlu 579

Db 602 AGGATTCAAACTTATTGTCAAATGAGTCCAATGTCCAGATTTTGGAACTCAACCCCTATGAA 661

QY 580 GlnTrpAla---IleGlnMetGluLysLysAlaAlaLysLysGlyAsnArgLysGluArg 598

Db 662 ACATGGGGCCCATTCCAATGGAAAAAAAGCCCTGCAAAAGAGGAAATCGCAAAGAACGT 721

QY 599 Val-CysAlaGluHisLeuArgLysTyrAsnGlu-AlaLeuGlnIle-AsnAspThrIle 617

Db 722 GTTTTGTGCAAAACATTTGAGGAAGTACATGAAGGCCCTACAAATTAAATGACACAATT 781

QY 618 ArgMet-IleAspAlaTyrThrHisLeu-GluThrPheTyrAsn-GluGluLysAsp-Ly 636

Db 782 TCAATGGATAAATGCGTATACTCATCTTTGAAACTTTTCTATATAATTGAGGAAAAGGATTAG 841

QY 636 sLysPheAla 639

Db 842 AAAGTTTGCC 851

RESULT 6

BU902097

LOCUS

DEFINITION BU902097 1174 bp mRNA linear EST 17-OCT-2002

AGENCOURT_10127740 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:6502757

5', mRNA sequence.

ACCESSION BU902097

VERSION BU902097.1 GI:24084010

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1174)

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Plate: LLAM14059 row: f column: 06

High quality sequence stop: 808.

Location/Qualifiers

1. .1174

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:6502757"

FEATURES

source

ORIGIN

Alignment Scores:

Pred. No.: 1.34e-123 Length: 1174

Score: 1257.00 Matches: 271

Percent Similarity: 87.81% Conservative: 10

Best Local Similarity: 84.69% Mismatches: 19

Query Match: 23.67% Indels: 20

DB: 5 Gaps: 1

US-09-515-363C-2 (1-1025) x BU902097 (1-1174)

1 MetSerAsnGlyTyrSerThrAspGluAsnPheArgTyrLeuIleSerCysPheArgAla 20

223 ATGTCGAATGGGTATTCCACAGACGAGAATTTCCGCTATCTCATCTCGTGTTCAGGGCC 282

21 ArgVallysMetTyrIleGlnValGluProValLeuAspTyrLeuThrPheLeuProAla 40

283 AGGGTGAAATGTACATCCAGGTGGAGCTGTGTGACTGACCTTTCTGCCTGCA 342

41 GluVallysGluGlnIleGlnArgThrValAlaThrSerGlyAsnMetGlnAlaValGlu 60

343 GAGGTGAAGGAGCAGATTCCAGAGGACAGTCGCCACCTCCGGGAACATGCAGGCAGTTGAA 402

61 LeuLeuLeuSerThrLeuGluLysGlyValTrpHisLeuGlyTrpThrArgGluPheVal 80

403 CTGCTGCTGAGCACCTTGGAGAAAGGAGCTGTGGACCTTGGTTGGACTCGGGAATCGTG 462

81 GluAlaLeuArgArgThrGlySerProLeuAlaAlaArgTyrMetAsnProGluLeuThr 100

463 GAGGCCCTCCGGAGAACCGGCAGCCCTCTGGCCCGCTACATGAACCCCTGAGCTCAG 522

101 AspLeuProSerProSerPheGluAsnAlaHisAspGluTyrLeuGlnLeuLeuAsnLeu 120

523 GACTTGCCCTCTCCATCGTTTGAGAACGCTCATGATGAATATCTCCAACCTGCTGAACCTC 582

121 LeuGlnProThrLeuValAspLysLeuLeuValArgAspValLeuAspLysCysMetGlu 140

583 CTTCAGCCCACTCTGGTGGACAAGCTTCTAGTTAGAGACGCTTGGATAAGTGATGGAG 642

141 GluGluLeuLeuThrIleGluAspArgAsnArgIleAlaAlaAlaGluAsnAsnGlyAsn 160

643 GAGGAACCTGTTGACAATTGAAGACAGAAACCGGATTGCTGCTGCAGAAACAATGGAAAT 702

161 GluSerGlyValArgGluLeuLeuLysArgIleValGlnLysGluAsnTrpPheSerAla 180

703 GAATCAGGTGTAAGAGAGACTACTAAAAAGGATTGTGCAGAAAGAAAACTGGTTCTCTGCA 762

181 PheLeuAsnValLeuArgGlnThrGlyAsnAsnGluLeuValGlnGluLeuThrGlySer 200

763 TTTCTGAATGTTCTCTCTCAACACAGGAAACAATGAACCTTGCTCAAGAGTTAACAGGCTCT 822

201 AspCysSerGluSerAsnAlaGluIleGluAsnLeuSerGlnValAspGlyProGlnVal 220

823 GATTGTCTCAGAAAGCAATGCAGAGATTGAGAATTTATCACAAGTTGATGGTCTCTCAAGTG 882

221 -GluGluGlnLeuLeuSerThrThr-ValGlnProAsnLeuGlu-LysGluValTrpGly 239

883 GGAAAAGCAACTCTTTCAACCACAGGTTTCAGCCAAATCTGGAAAAAGGAGGTTCTGGGGC 942

240 MetGlu-AsnAsnSerSer-GluSerSerPheAlaAspSerSerValValSerGluSera 259

943 ATGGAAAAATAACTCATCAAAAATCATCTTTTGCAGATCTCTGAAAGTTCCAGAACCAA 1002

259 spThrSerLeuAla-GluGlySerValSerCysLeu-AspGluSer-LeuGlyHisAsns 278

1003 AACAAAGTTTGGGCAAAAGGAAATGGCAACTGCTTAAATGAAAGTCTTGGGACATAACA 1062

278 erAsnMetGlySerAsp-----SerGlyT 286

1063 ACAACATGGGCAGGATTTCCGGCCCCCGGGCAAGGATTTCAAATAAAAAAAGGGG 1122

286 hrMetGlySerAspSerAspGluGluAsnValAlaAlaArgAla 300

1123 CCACCAAGAAAAATTCCTCCGGGAACAAAAAATTCAGCTTGGGCC 1166

RESULT 7

BM467983

LOCUS BM467983 1115 bp mRNA linear EST 05-FEB-2002

DEFINITION AGENCOURT_6437921 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5532884 5', mRNA sequence.

ACCESSION BM467983

VERSION BM467983.1 GI:18517025

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1115)

NIH-MGC http://mgc.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: ATCC

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov

Plate: LLAM12216 row: n column: 21

High quality sequence stop: 690.

Location/Qualifiers

1.1115

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5532884"

/tissue_type="leiomyosarcoma"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH_MGC_71"

/note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2.1 kb."

ORIGIN

Alignment Scores:

Pred. No.: 1.91e-119 Length: 1115

Score: 1218.00 Matches: 228

Percent Similarity: 99.13% Conservative: 0

Best Local Similarity: 99.13% Mismatches: 2

Query Match: 22.93% Indels: 0

DB: 4 Gaps: 0

US-09-515-363C-2 (1-1025) x BM467983 (1-1115)

796 LeuAspIleLysGluCysAsnIleValIleArgTyrGlyLeuValThrAsnGluIleAla 815

1 TTGGATATTAAAGAAATGTAACATTGTTATCCGTTATGGTCTCGTCACCAATGAATAGCC 60

816 MetValGlnAlaArgGlyArgAlaArgAlaAspGluSerThrTyrValLeuValAlaHis 835

61 ATGGTCCAGGCCCGTGTGCGAGCCAGAGCTGATGAGAGCACCTACGTCCTGGTGTCTCAC 120

836 SerGlySerGlyValIleGluHisGluThrValAsnAspPheArgGluLysMetMetTyr 855

121 AGTGGTTCCAGAGTTATCGAACCGTGAGACAGTAAATGATTTCCGAGAGAGATGATGTAT 180

856 LysAlaIleHisCysValGlnAsnMetLysProGluGluTyrAlaHisLysIleLeuGlu 875

```
Db      ||||| 181 AAAGCTATACATTGTTCAAAATATGAAACAGAGGAGTATGCTCATAGATTTTGGAA 240
Qy      ||||| 876 LeuGlnMetGlnSerIleMetGluLysLysMetLysThrLysArgAsnIleAlaLysHis 895
Db      ||||| 241 TTACAGATGCAAAGTATAATGGAAGAAATGAAACCAAGAGAAATATTGCCAAGCAT 300
Qy      ||||| 896 TyrLysAsnAsnProSerLeuIleThrPheLeuCysLysAsnCysSerValLeuAlaCys 915
Db      ||||| 301 TACAAGAATAACCCCATCACTAATACTTCTCTTGCAAAACCTGCAGTGTGCTAGCCCTGT 360
Qy      ||||| 916 SerGlyGluAspIleHisValIleGluLysMetHisHisValAsnMetThrProGluPhe 935
Db      ||||| 361 TCTGGGAAGATATCCATGTAATTGAGAAATGCATCACGTCAATATGACCCAGAAATTC 420
Qy      ||||| 936 LysGluLeuTyrIleValArgGluAsnLysAlaLeuGlnLysLysCysAlaAspTyrGln 955
Db      ||||| 421 AAGGAACTTTTACATTGTAAGAGAAAAACAAACACTGCAAAAGAGTGTGCCGACTATCAA 480
Qy      ||||| 956 IleAsnGlyGluIleIleCysLysCysGlyGlnAlaTrpGlyThrMetMetValHisLys 975
Db      ||||| 481 ATAAATGGTGAATCATCTGCAATGTGCCAGGCTTGGGAACAATGATGGTGCACAAA 540
Qy      ||||| 976 GlyLeuAspLeuProCysLeuLysIleArgAsnPheValValPheLysAsnAsnSer 995
Db      ||||| 541 GGCTTAGATTTGCCCTTGCTCAAAATAAGGAATTTGTAGTGGTTTCAAAAATAATCA 600
Qy      ||||| 996 ThrLysLysGlnTyrLysLysTrpValGluLeuProIleThrPheProAsnLeuAspTyr 1015
Db      ||||| 601 ACAAGAAACAATACAAAAAGTGGGTAGAAATACCTATCACATTTCCCAATCTTGACTAT 660
Qy      ||||| 1016 SerGluCysCysLeuPheSerAspGluAsp 1025
Db      ||||| 661 TCAGAAATGCTGTTTATTAGTAGATGAGGAT 690
```

```
RESULT 8
BF337464
LOCUS      BF337464
DEFINITION 602035195F1 NCI_CGAP_Brn64 Homo sapiens cDNA clone IMAGE:4183126
5', mRNA sequence.
ACCESSION  BF337464
VERSION    BF337464.1 GI:11283715
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1035)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: David N. Louis, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9498 row: n column: 23
High quality sequence start: 5
High quality sequence stop: 695.
FEATURES             Location/Qualifiers
     source           1..1035
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4183126"
                     /tissue_type="glioblastoma with EGFR amplification"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NCI_CGAP_Brn64"
                     /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
```

```
Site_2: Sali; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.57 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
ORIGIN
Alignment Scores:      2.92e-118      Length:      1035
Pred. No.:            1206.50      Matches:      264
Score:                84.04%      Conservative: 15
Best Local Similarity: 79.52%      Mismatches:   44
Query Match:         22.72%      Indels:      12
DB:                   2           Gaps:           5
US-09-515-363C-2 (1-1025) x BF337464 (1-1035)
Qy      295 AsnValAlaAlaArgAlaSerProGluProGluLeuGlnLeuArgProTyrGlnMetGlu 314
Db      6 AATGTGGCAGCAAGAGCATCCCCGGAGCCAGAACTCCAGCTCAGGCCTTACCAGATGGAA 65
Qy      315 ValAlaGlnProAlaLeuGluGlyLysAsnIleIleCysLeuProThrGlySerGly 334
Db      66 GTTGCCCGAGCCAGCCTTGGAAGGGAAGAATATCATCATCTGCCTCCCTACAGGGAGTGA 125
Qy      335 LysThrArgValAlaValTyrIleAlaLysAspHisLeuAspLysLysLysLysAlaSer 354
Db      126 AAAACCAGAGTGGCTGTTTACATTGCCAAGGATCACCTTAGACAAGAAAAAGCATCT 185
Qy      355 GluProGlyLysValIleValLeuValAsnLysValLeuValGluGlnLeuPheArg 374
Db      186 GAGCCTGGAAAAAGTTATAGTTCTTGTCAATAAGGTACTGCTAGTTGAACAGCTCTTCGCG 245
Qy      375 LysGluPheGlnProPheLeuLysLysTrpTyrArgValIleGlyLeuSerGlyAspThr 394
Db      246 AAGGAGTTCCAACCATTTTGAAGAAATGGTATCGTGTATTGGATTAAAGTGTGTATACC 305
Qy      395 GlnLeuLysIleSerPheProGluValValLysSerCysAspIleIleIleSerThrAla 414
Db      306 CAACTGAAATATCATTTCCAGAAAGTTGTCAAGTCTCTGTATATTATTATCAGTACAGCT 365
Qy      415 GlnIleLeuGluAsnSerLeuLeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeu 434
Db      366 CAAATCCTTGAAAAACTCCCTCTTAAACTTGGAAATGGAGAAGATGCTGGTGTTCATTG 425
Qy      435 SerAspPheSerLeuIleIleAspGluCysHisThrAsnLysGluAlaValTyr 454
Db      426 TCAGACTTTTCCCTCATTTATCATTGATGAATGTATCATCACACCAACAAAGNAGCAGTGTAT 485
Qy      455 AsnAsnIleMetArgHisTyrLeuMetGlnLysLeuLysAsnAsnArgLeuLysLysGlu 474
Db      486 AATAACATCATGAGGCATTATTTGTATGCAGAAAGTTGAAAAACAATAGACTCAAGAAAGAA 545
Qy      475 AsnLysProValIleProLeuProGlnIleLeuGlyLeuThrAlaSerProGlyValGly 494
Db      546 AACAAACCAGTGATTTCCCTTCTCAGATACTGGGACTAACAGCTTCACCTGGTGTCCGA 605
Qy      495 GlyAlaThrLysGlnAlaLysAlaGluGluHisIleLeuLysLeuCysAlaAsnLeuAsp 514
Db      606 GGGGCCACGAAGCAAGCCAAAGCTGAAGAACACATTTTAAAACTATGTGCCAATCT-GAT 664
Qy      515 AlaPheThrIleLysThrValLysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGlu 534
Db      665 GCATTTACTATTAAAACGTGTGTAGAAAAACCTTGATCACTGGGAAAAACCAATACAGAG 724
Qy      535 ProCysLysLysPheAlaIleAlaAspAlaThrArgGluAspProPheLysGluLysLeu 554
Db      725 CCATGCGAGAAGTTGGCGCT-GCGATGCGAC-CGAGAAGATCCCTTTAAGGGAAAC--- 779
Qy      555 LeuGluIleMetThrArgIleGlnThrTyrCysGlnMetSerPro-MetSerAspPheG1 574
Db      780 ---TCTGAAATAGGCAGGGTTCCTTGGTGCAGT---AGTCCAGGTCCGCA---TTGGG 830
Qy      574 yThrGlnProTyrGluGlnTrpAlaIleGlnMetGluLysLysAlaAlaLysLysGlyAs 594
```

D _b	831	TAGCCACCTATTAAA-----GGGGCCTCAACGGGGGAAAAAGGGGGCTTAAGAACGACAA	884
Q _y	594	nArgLysGluArgValCysAlaGluHisLeuArgLysTyrAsnGluAlaLeuGlnIleAs	614
D _b	885	ACGGTTGGSCGACCTTGGGGCGAAACAAAGGCGCACACAGAGACACACAGAAGAGGGCGACACAA	944

Qy 614 nAspThrIleargMetIleAspAlaTyrThrHis 625
 |::: |::: |||||
Db 945 CGAACACAGCAAA-----GACAGAACGACACAC 972

RESULT 9	
BF983236	
LOCUS	729 bp mRNA
DEFINITION	BF983236 602305873F1 NIH_MGC_88 Homo sapiens cDNA clone IMAGE:4397083 5', mRNA sequence.

ORIGIN

Alignment Scores:		
Pred. No.:		729
Score:	3.55e-110	238
Percent Similarity:	1129.00	2
Best Local Similarity:	98.36%	Conservative:
Query Match:	97.54%	Mismatches: 4
	21.26%	Indels: 3
DB:	4	Gaps: 0

US-09-515-363C-2 (1-1025) x BF983236 (1-729)

[illegible]

Qy 160 AsnGluSerGlyValArgGluLeuLeuLysArgIleValGlnLysGluAsnTrpPheSer 179
|||
Db 61 AATGAATCAGGTGTAAGAGAGCTACTAAAAAGGATTGTGCAGAAAGAAAA CTGGTTCTCT 120

Qy 180 AlaphelLeuAsnValLeuArgGlnThrGlyAsnAsnGluLeuValGlnLeuThrGly 199

Db	121	GCATTTCTGAATGTTCTTCGTCAAAACAGGAAACAATGAACCTTGTCCAAAGAGTTAACACAGGC	180
Qy	200	SerAspCysSerGluSerAsnAlaGluIleGluAsnLeuSerGlnValAspGlyProGln	219
Db	181	TCTGATTGCTCAGAAAGCAATGCAGAGATTGAGAAATTTATCACAAAGTTGATGGTCTCTCAA	240
Qy	220	ValGluGluGlnLeuLeuSerThrThrValGlnProAsnLeuGluLysGluValTrpGly	239
Db	241	GTGGAAGAGACAACTTCTTTCAACACACAGTTTCAGCCAAATCTGGAGAAGGAGGTCTGGGGC	300
Qy	240	MetGluAsnAsnSerSerGluSerSerPheAlaAspSerSerValValSerGluSerAsp	259
Db	301	ATGGAGAATAACTCATCAGAATCATCTTTTGCAGATCTTCTGTAGTTTCAGAAATCAGAC	360
Qy	260	ThrSerLeuAlaGluGlySerValSerCysLeuAspGluSerLeuGlyHisAsnSerAsn	279
Db	361	ACAAGTTTGGCAGAAAGGAAGTGTCTGCTGCTTATGATAAAGTCTTGGACATAACAGCAAC	420
Qy	280	MetGlySerAspSerGlyThrMetGlySerAspSerAspGluGluAsnValAlaAlaArg	299
Db	421	ATGGGCAGTGATTTCAGGCACCATGGGAAGTGATTTCAGATGAAGAGAAATGTGGCAGCAAGA	480
Qy	300	AlaSerProGluProGluLeuGlnLeuArgProTyrGlnMetGluValAlaGlnProAla	319
Db	481	GCATCCCGGAGCCAGAACTCCAGCTCAGGCCCTTACCAAATGGAAGTTGCCACAGCCAGC-	539
Qy	320	LeuGluGlyLysAsnIleIleCysLeuProThrGlySerGlyLysThrArgValAla	339
Db	540	TTGGAAGGGAAGAATATCATCATCTGTCTC-CCTACAGGGAGTGGAAAAACACAGATGGCT	598
Qy	340	ValTyrIleAlaLysAspHisLeuAspLysLysLysAlaSerGluProGlyLysVal	359
Db	599	GTTTACATTGCCAAGGATCATTAGACAG-AAGAAAAAAGCATCTGAGCCCTGGAAAAAGTT	657
Qy	360	IleValLeuValAsnLysValLeuLeuValGluGlnLeuPheArgLysGluPheGlnPro	379
Db	658	ATAGTCTTGTGCATAAGGTACTGCTAGTTGAACAGCTCTTCCGAAAGGAGTTCCCAACCA	717
Qy	380	PheLeuLysLys 383	
Db	718	TTTTTGGAGAAG 729	

RESULT 10

BY720783

LOCUS

DEFINITION

RECEIVED

ACCESSION
NUMBER

KEYWORDS:

SOURCE

ORGANISM

REFERENCE

AUTHORS

Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y., Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I., Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P., Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M., Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K., Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K., Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S., Rogers,J., Birney,E. and Hayashizaki,Y.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
22354683
12466851
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@sc.riken.jp, URL:http://genome.gsc.riken.jp/
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hashizume,W., Hayashida,K., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Koya,S., Miyazaki,A., Murata,M., Nakamura,M., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Saito,R., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami,M., Takeda,Y., Waki,K., Watahiki,A., Muramatsu,M. and Hayashizaki,Y.
Direct Submission
Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site (http://genome.gsc.riken.go.jp) for further details.

FEATURES

source
1. .972
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="9130009C22"
/sex="male"
/tissue type="cecum"
/dev_stage="adult"
/lab_host="DH108"
/clone_lib="RIKEN full-length enriched, adult male cecum"
/notes="Site 1: Sali; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5'
GAGAGAGAGGATCCAGAGCTCTTTTCTTTTCTTTT 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 185.0. Second strand cDNA was prepared with the primer adapter of

sequence [5' GAGAGAGAGATTCTCGAGTTAATAATTAATCCCCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: Sali; 3' end: BamHI"

ORIGIN

Alignment Scores:
Pred. No.: 3.17e-109 Length: 972
Score: 1122.00 Matches: 215
Percent Similarity: 92.62% Conservative: 11
Best Local Similarity: 88.11% Mismatches: 18
Query Match: 21.13% Indels: 0
DB: 6 Gaps: 0
US-09-515-363C-2 (1-1025) x BY720783 (1-972)
QY 782 LysIleAsnLeuIleAlaThrValAlaGluGlyLeuAspIleLysGluCys 801
Db 1 GAAATAAATCTGCTTATCGCTACGACGGTGGCAGAGGAGCTGGATATCAAAGAGTGC 60
QY 802 AsnIleValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAlaArgGly 821
Db 61 AATATTGTTATTCGTTATGGCCTTGTTCACGAACGAGATAGCCATGGTCCAGGCCCGGGT 120
QY 822 ArgAlaArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGlyValIle 841
Db 121 CGAGCCAGAGCTGATGAAAGCACGTATGTCCTGGTCACCAGCAGTGGCTCAGGAGTTACC 180
QY 842 GluHisGluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHisCysVal 861
Db 181 GAACGGGAGATTGTTAATGATTTCCGAGAGAAGATGATGTATATAAGCTATTAAACCGTGT 240
QY 862 GlnAsnMetLysProGluGluTyrAlaHisLysIleLeuGluLeuGlnMetGlnSerIle 881
Db 241 CAAACATGAAACCAGAGGAGTATGCACATAAGATTTTGGAAATTGCAGGTGCAAAAGTATC 300
QY 882 MetGluLysLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAsnProSer 901
Db 301 CTGGAAAAGAAAATGAAAGTCGAAAGAAGCATTTGCAAAAGCAATACAAACGACATCCATCG 360
QY 902 LeuIleThrPheLeuCysLysAsnCysSerValLeuAlaCysSerGlyGluAspIleHis 921
Db 361 TTAATAACACTTCTCTGCAAAAATTTAGCATGCTGGTCTGCTCGGAGAAAACATCCAT 420
QY 922 ValIleGluLysMetHisHisValAsnMetThrProGluPheLysGluLeuTyrIleVal 941
Db 421 GTCATTGAGAAGATGCATCATGTCAATATGACACCAGAAATTCAGGGACTCTACATTGTA 480
QY 942 ArgGluAsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluIleIle 961
Db 481 AGAGAAAACAAAGCACTGCAAAAGAAATTTGCTGATTATCAGACCAATGGAGAGATTATC 540
QY 962 CysLysCysGlyGlnAlaTrpGlyThrMetMetValHisLysGlyLeuAspLeuProCys 981
Db 541 TGCAAGTGTGGCCAGGCTTGGGGAACAATGATGGTGCAAAAGGTTTAGATTTGCCTGT 600
QY 982 LeuLysIleArgAsnPheValValPheLysAsnAsnSerThrLysLysGlnTyrLys 1001
Db 601 CTTAAATAAGGAATTTTGTAGTCAATTTCAAAAATAACTCACCGAAGAAACAGTACAAG 660
QY 1002 LysTrpValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCysCysLeuPhe 1021
Db 661 AAGTGGGTGGAATTGCCTATCAGATTTCTGTGATCTTGACTACTCAGAATACTGCTTGAT 720
QY 1022 SerAspGluAsp 1025
Db 721 AGTGATGAAGAT 732

RESULT 11

AK018602
LOCUS AK018602 1197 bp mRNA linear HTC 03-APR-2004
DEFINITION Mus musculus adult male cecum cDNA, RIKEN full-length enriched

library, clone:9130009C22 product:similar to MELANOMA
DIFFERENTIATION ASSOCIATED PROTEIN-5 [Homo sapiens], full insert
sequence.
AK018602
VERSION AK018602.1 GI:12858393
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159
2
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159
3
Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861
4
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 1197)
Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Arai,A., Aono,H.,
Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Fukunishi,Y.,
Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K.,
Hiraoka,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Izawa,M.,
Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K.,
Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C.,
Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D.,
Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y.,
Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T.,
Tejima,Y., Toya,T., Yamamura,T., Yasunishi,A., Yoshida,K.,
Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
Direct Submission
Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp,
URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)
please visit our web site (http://genome.gsc.riken.jp/) for further
details.
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken

Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues. First strand cDNA was primed with a primer
[5' GAGAGAGAGAGGATCAAGAGCTCTTTTCTTTTCTTTTNN 3'], cDNA was
prepared by using trehalose thermo-activated reverse transcriptase
and subsequently enriched for full-length by cap-trapper. cDNA went
through one round of normalization to Rot = 10.0 and subtraction to
Rot = 185.2. Second strand cDNA was prepared with the primer
adapter of sequence [5'
GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCC 3']. cDNA was cleaved
with BamHI and XhoI. Vector: a modified pBluescript KS(+) after
bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3'
end: BamHI. Host: DH10B.
FEATURES
source
1..1197
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM DB:9130009C22"
/db_xref="taxon:10090"
/clone="9130009C22"
/sex="male"
/tissue_type="cecum"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
CDS
<1..737
/note="unnamed protein product; putative
similar to MELANOMA DIFFERENTIATION ASSOCIATED PROTEIN-5
[Homo sapiens] (SPTR|Q9H3G6, evidence: FASTY, 75.6%ID,
67.6%length, match=2079)"
/codon_start=3
/protein_id="BAB31303.2"
/db_xref="GI:26384184"
/translation="EINLLIATTVAEGLDIKECNIVIRYGLVTNEIAMVQARGRARA
DESTYVLVTSGSGVTEREIVNDFREKMYKAINRVQNMKPEYAHKILELQVQSILE
KMKVERSIKQYNDNPSLITLLCKNSMLVCSGENIHVIEKMHVNMTPFEKGLYIV
RENKALQKFADYQTNGEIIICKGOAWGTMVHKGLDLPCLKIRNFVNFKNNSPKKQ
YKKWELPIRFPDDL DYSECLYSD E"
1180..1185
/note="putative"
1197
/note="putative"
ORIGIN
Alignment Scores:
Pred. No.: 4.41e-109 Length: 1197
Score: 1122.00 Matches: 215
Percent Similarity: 92.62% Conservative: 11
Best Local Similarity: 88.11% Mismatches: 18
Query Match: 21.13% Indels: 0
DB: 3 Gaps: 0
US-09-515-363C-2 (1-1025) x AK018602 (1-1197)
QY 782 LysIleAsnLeuLeuIleAlaThrThrValAlaGluGlyLeuAspIleLysGluCys 801
:::|||||
Db 3 GAAATAAATCTGCTTATCGCTACGACGGTGGCAGAGGAAGGCCTGGATATCAAAGAGTGC 62
QY 802 AsnIleValIleArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAlaArgGly 821
|||||
Db 63 AATATTGTTATTTCGTTATGCTTGTTCACGAACGAGATAGCCATGGTCCAGGCCCGGGT 122
QY 822 ArgAlaArgAlaAspGluSerThrTyrValLeuValAlaHisSerGlySerGlyValIle 841
|||||
Db 123 CGAGCCAGAGCTGATGAAGACACGATATGTCCTGTGCACCAGCAGTGGCTCAGGAGTTACC 182
QY 842 GluHisGluThrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHisCysVal 861
|||
Db 183 GAACGGGAGATTGTTAATGATTTCCGAGAGAGAGATGATGATATAAGCTATTAAACCGTGT 242
QY 862 GlnAsnMetLysProGluGluTyrAlaHisIleLeuGluLeuGlnMetGlnSerIle 881
|||||
Db 243 CAAACATGAACACCAGAGAGTATGCACATAAGATTTTGAATTGCAGGTGCAAGTATC 302

QY 882 MetGluLysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAsnProSer 901
Db 303 CTGGAAGAAATGAAAGTCGAAAGAGCATTCGAAAGCAATACACGACAATCCATCG 362
QY 902 LeuIleThrPheLeuCysLysAsnCysSerValLeuAlaCysSerGlyGluAspIleHis 921
Db 363 TTAATAACACTTCTCTGCAAAAATTGTAGCATGCTGGTCTGCTCGGAGAAACATCCAT 422
QY 922 ValIleGluLysMetHisHisValAsnMetThrProGluPheLysGluLeuTyrIleVal 941
Db 423 GTCATTGAGAAGATGCATCATGTCAATATGACACCAGAATTCAAGGGACTCTACATTGTA 482
QY 942 ArgGluAsnLysAlaLeuGlnLysLysCysAlaAspTyrGlnIleAsnGlyGluIleIle 961
Db 483 AGAGAAAACAAAGCACTGCAAAAGAAATTGCTGATTATCAGACCAATGGAGAGATTATC 542
QY 962 CysLysCysGlyGlnAlaTrpGlyThrMetMetValHisLysGlyLeuAspLeuProCys 981
Db 543 TGCAAGTGTGGCAGGCTTGGGGAACAATGATGGTGCACAAAGGTTTAGATTTCCTTGT 602
QY 982 LeuLysIleArgAsnPheValValPheLysAsnAsnSerThrLysLysGlnTyrLys 1001
Db 603 CTTAAATAAGGAATTTGTAGTCAATTTCAAAAATAACTCACCGAAGAACAGTACAAG 662
QY 1002 LysTrpValGluLeuProIleThrPheProAsnLeuAspTyrSerGluCysCysLeuPhe 1021
Db 663 AAGTGGTGAATTGCCTATCAGATTTCTGTATCTTGACTACTCAGAATACTGCTTGAT 722
QY 1022 SerAspGluasp 1025
Db 723 AGTGATGAAGAT 734

RESULT 12
BG7411146
LOCUS
DEFINITION
602631817F1 NCI_CGAP_Skn3 Homo sapiens cDNA clone image:4776869 5',
mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
BG7411146
BG7411146.1 GI:14051799
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
1 (bases 1 to 827)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov

Tissue Procurement: James Cleaver, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA
Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM10630 row: b column: 06
High quality sequence stop: 763.

FEATURES
source
Location/Qualifiers
1..827

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4776869"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI CGAP Skn3"
/note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

ORIGIN

Alignment Scores:
Pred. No.: 9.3e-107 Length: 827
Score: 1098.00 Matches: 255
Percent Similarity: 91.17% Conservative: 3
Best Local Similarity: 90.11% Mismatches: 13
Query Match: 20.67% Indels: 15
DB: 4 Gaps: 0
US-09-515-363C-2 (1-1025) x BG7411146 (1-827)
QY 87 GlySerProLeuAlaAlaArgTyrMetAsnProGluLeuThrAspLeuProSerProSer 106
Db 2 GGCAGCCCTCTGGCCGC-CGCTACATGAACCCCTGAGCTCAGGACTTGCCCTCTCCATCG 60
QY 107 PheGluAsnAlaHisAspGluTyrLeuGlnLeuLeuAsnLeuGln-ProThrLeuVa 126
Db 61 TTTGAGAACGCTCATGATGAATATCTCCAACCTGCTGAACCTCCTTCAGTCCCACCTGGT 120
QY 126 lAspLysLeuLeuValArgAspValLeuAspLysCysMetGluGluLeuThrIle 146
Db 121 GGACAAGCTTCTAGTAGAGACGCTTGGATAAGTGCATGAGGAGGAACCTGTTGACAAAT 180
QY 146 eGluAspArgAsnArgIleAlaAlaAlaGluAsnAsnGlyAsnGluSerGlyValArgG1 166
Db 181 TGAAGACAGAAACCGGATTGCTGCTGCAGAAAACAATGGAATGAATCAGGTGTAAGAGA 240
QY 166 uLeuLeuLysArgIleValGlnLysGluAsnTrpPheSerAlaPheLeuAsnValLeuAr 186
Db 241 GCTACTAAAAAGGATTGTGCAGAAAGAAAACCTGGTTCTCTGCATTTCTGAATGTTCTTCG 300
QY 186 gGlnThrGlyAsnAsnGluLeuValGlnGluLeuThrGlySerAsp--CysSerGluSer 205
Db 301 TCAAACACAGGAACAATGAACCTTGCCACGAGTTAACAGGCTCTGATTGGCTCAGAAAAAGC 360
QY 206 AsnAlaGluIleGluAsnLeuSerGlnValAspGlyProGlnValGluGlnLeuLeu 225
Db 361 AATGCAGAGATTGAGAAATTTATCACAAGTTGATGGTCTCTCAAGTGAAGAGCAACTTCTT 420
QY 226 SerThrThrValGlnProAsnLeuGlu-LysGluValTrpGlyMetGluAsnAsnSerSe 245
Db 421 TCAACCACAGTTCAGCCAAATCTGGACGAAGGAGGTCTGGGGCATGGAGAAATAACTCATC 480
QY 245 rGluSerSerPheAla-AspSerSerValValSerGluSerAspThrSerLeuAla--G1 264
Db 481 AGAATCATCTTTTGCACGATTCTCTGTAGTTTCAGAAATCAGACACAAGTTTGGCACGCA 540
QY 264 uGlySerValSerCysLeuAspGluSerLeuGly-HisAsnSerAsnMetGlySerAsps 284
Db 541 AGGAAGTGTGAGTGTCTTAGATGAAAGTCTTTGGACCATAACAGCAACATGGGCAGTGATT 600
QY 284 erGlyThrMetGlySerAspSerAspGlu-GluAsnValAlaAlaArgAlaSerProGlu 303
Db 601 CAGGCACCATGGGAAGTGATTTCAGATGAACGAGAATGTGGCAGCAAGAGCATCCCCGGAG 660
QY 304 ProGluLeuGlnLeuArgProTyrGlnMet--GluValAlaGlnProAlaLeuGluGlyL 323
Db 661 CCAGAACTCCAGTCCAGGCCCTTACCAATGGCACAGTTGCCAGCCAGC-TTGGCAGGGA 719
QY 323 ysAsnIleIleIleCysLeuProThrGlySerGlyLysThrArgValAlaValTyrIle- 342
Db 720 AGACTATCATCATCTGCTC-CCTACAGGGAGTGGAAAAACAGAAAGTGGCTGATTACATTA 778
QY 343 AlaLysAspHisLeuAspLysLysLysAlaSerGluProGly 357
Db 779 GCCACGGATCACTTAGACAAGAAGAAACAAGCATCTGAGCCTGGA 823
RESULT 13
BF686405
LOCUS
DEFINITION
602143786F1 NIH_MGC_46 Homo sapiens cDNA clone image:4304805 5',
mRNA sequence.
ACCESSION
BF686405

ORIGIN

Alignment Scores: 7.95e-105 Length: 1041
Pred. No.: 1081.50 Matches: 232
Score: 77.04% Conservative: 13
Percent Similarity: 72.96% Mismatches: 40
Best Local Similarity: 20.36% Indels: 34
Query Match: 2 Gaps: 4
DB:

US-09-515-363C-2 (1-1025) x BE882040 (1-1041)

QY 586 GluLysLysAlaAlaLysLysGlyAsnArgLysGluArgValCysAlaGluHisLeuArg 605
Db 2 GAAAAAAGCTGCAAAAGAGGAAATCGCAAGAACGTTGTTGTCAGAACATTTGAGG 61
QY 606 LysTyrAsnGluAlaLeuGlnIle-AsnAspThrIleArgMetIleAspAlaTyrThrHi 625
Db 62 AAGTACAAGGAGGCCCTACAAATTAAGTGACACAAATTCGAATGATAGATCGGTACTCA 121
QY 625 sLeuGluThrPheTyrAsnGluGluLysAspLysLysPheAlaValIleGluAspAspSe 645
Db 122 TCTTGAACCTTCTATAATGAAGAGAGAAAGATGAAGAGTTTGCAGTCATAGAAGATGATAG 181
QY 645 rAspGluGlyGlyAspAspGluTyrCysAspGlyAspGluAspGluAspLeuLysLy 665
Db 182 TGATGAGGGTGGTGATGATGAGTATTGTCATGATGAAGATGAGGATGATTTAAAGAA 241
QY 665 sProLeuLysLeuAspGluThrAspArgPheLeuMetThrLeuPhePheGluAsnAsnLy 685
Db 242 ACCTTTGAAACTGGATGAACAGATAGATTCTCATGACTTTATTTTGGAAAACAATAA 301
QY 685 sMetLeuLysArgLeuAlaGluAsnProGluTyrGluAsnGluLysLeuThrLysLeuAr 705
Db 302 AATGTTGAAAAGGCTGGCTGAAAACCCAGCAATATGAAAATGAAAAGCTGACCAAAATTAAG 361
QY 705 gAsnThrIleMetGluGlnTyrThrArgThrGluGluSerAlaArgGlyIleIlePheTh 725
Db 362 AAATACCATAATGGAGCAATATACTAGGACTGAGGAATCAGCACGAGGAATAATCTTTAC 421
QY 725 rLysThrArgGlnSerAlaTyrAlaLeuSerGlnTrpIleThrGluAsnGluLysPheAl 745
Db 422 AAAAAACACGACAGAGTGCATATCGCTTTCCAGTGGATTACTGAAAATGAAAAATTTGC 481
QY 745 aGluValGlyValLysAlaHisHisLeuIleGlyAlaGlyHisSerSerGluPheLysPr 765
Db 482 TGAAGTAGGAGTCAAAGCCCAACCATCTGATTGGAGCTGGACACAGCAGTGAGTTCAAACC 541
QY 765 oMetThrGlnAsnGluGlnLysGluValIleSerLysPheArgThrGlyLys-IleAsnL 785
Db 542 CATGACACAGAAATGAACAAAGAAAGTATTAGTAAATTTCCGACTGGCAACAATAATC 601
QY 785 euLeuIleAlaThrThrValAlaGluGluGlyLeuAspIleLysGluCysAsnIleValI 805
Db 602 TGCTTATCGCTACACAGTGGCAGAGAGAGGTCCGGATATTCCAGCTGTACC---ACGC 658
QY 805 leArgTyrGlyLeuValThrAsnGluIleAlaMetValGlnAlaArgGlyArgAlaArgA 825
Db 659 TATCCGAGGCCCCAGCACCAATGAATAGGCCATGGC-CAAGCCGGGCCCCAAGCCAAAC 717
QY 825 laAspGluSerThrTyrValLeuValAlaHisSerGlySerGlyValIleGluHisGluT 845
Db 718 TACTGAAAAGACACCAAAAGACGGCGTGGCCACAAAGGTCA-----CAGCAGA 765
QY 845 hrValAsnAspPheArgGluLysMetMetTyrLysAlaIleHisCysValGlnAsnMetL 865
Db 766 CAACCAAC-----GCCACACAGCCAGGACATCCCAGA 798
QY 865 ysProGluGluTyrAlaHisLysIleLeuGluLeuGlnMetGlnSerIleMetGluLysL 885
Db 799 AAACACAGCATGCACACCACCCCAACATGCGG----- 830
QY 885 ysMetLysThrLysArgAsnIleAlaLysHisTyrLysAsnAsnProSer 901

Db 831 -----GCCACACACTATCAAAACCCAGCGAGC 857
RESULT 15
BX492926
LOCUS
DEFINITION BX492926 627 bp mRNA linear EST 04-SEP-2003
DKFZp781C0628 r1 781 (synonym: hlcc4) Homo sapiens cDNA clone
ACCESSION DKFZp781C0628 5', mRNA sequence.
VERSION BX492926
KEYWORDS BX492926.1 GI:32004887
SOURCE EST.
ORGANISM Homo sapiens (human)
REFERENCE Homo sapiens
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 627)
Wambutt,R., Heubner,D., Mewes,H.W., Weil,B., Amid,C., Osanger,A.,
Fobo,G., Han,M. and Wiemann,S.
EST (Wambutt,R., Heubner,D., Mewes,H.W., Weil,B., Amid,C., et al.)
Unpublished (2003)
Contact: MIPS
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing
consortium of the German Genome Project.
No s1 sequence available.
This clone (DKFZp781C0628) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
Location/Qualifiers
1..627
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp781C0628"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="781 (synonym: hlcc4)"
/note="Vector: pSport1_Sfi; Site_1: SfiIA; Site_2: SfiIB;
cDNA-collection"

ORIGIN

Alignment Scores: 8.43e-105 Length: 627
Pred. No.: 1078.00 Matches: 208
Score: 100.00% Conservative: 1
Percent Similarity: 99.52% Mismatches: 0
Best Local Similarity: 20.30% Indels: 0
Query Match: 5 Gaps: 0
DB:
US-09-515-363C-2 (1-1025) x BX492926 (1-627)
QY 402 GluValValLysSerCysAspIleIleIleSerThrAlaGlnIleLeuGluAsnSerLeu 421
Db 1 GAAGTTGTCAGTCTCTGTGATATTATTATCAGTACAGCTCAAATCCTTGAAAACCTCCCTC 60
QY 422 LeuAsnLeuGluAsnGlyGluAspAlaGlyValGlnLeuSerAspPheSerLeuIleIle 441
Db 61 TTAACCTTGAAAAATGGAGAGAGTCTGGTGTTCATATTGTCAGACTTTTCCCTCATATATC 120
QY 442 IleAspGluCysHisHisThrAsnLysGluAlaValTyrAsnAsnIleMetArgHisTyr 461
Db 121 ATTGATGAATGTCATCACACCAACAAAGAGCAGTGATATAATAACATCATGAGGCATTAT 180
QY 462 LeuMetGlnLysLeuLysAsnArgLeuLysLysGluAsnLysProValIleProIle 481
Db 181 TTGATGCAGAAGTTGAAAAACAATAGACTCAAGAAAGAAAAACAACACAGTGATCCCTT 240
QY 482 ProGlnIleLeuGlyLeuThrAlaSerProGlyValGlyAlaThrLysGlnAlaLys 501

```
Db      241 CCTCAGATACTGGGACTAACAGCTTCACCTGGTGTGGAGGGGCCACGAAGCAAGCCAAA 300
QY      502 AlaGluGluHisIleLeuLysLeuCysAlaAsnLeuAspAlaPheThrIleLysThrVal 521
Db      301 GCTGAAGAACACACATTTTAAACTATGTGCCAATCTTGATGCATTTACTATTAAACTGTT 360
QY      522 LysGluAsnLeuAspGlnLeuLysAsnGlnIleGlnGluProCysLysLysPheAlaIle 541
Db      361 AAAGAAAACCTTGATCAACTGAATAAACCAATACAGGAGCCATGCAAGAGTTTGCCATT 420
QY      542 AlaAspAlaThrArgGluAspProPheLysGluLysLeuGluIleMetThrArgIle 561
Db      421 GCAGATGCAACACAGAGAGATCCATTAAAGAGAACTTCTAGAAATAATGACAAGGATT 480
QY      562 GlnThrTyrCysGlnMetSerProMetSerAspPheGlyThrGlnProTyrGluGlnTrp 581
Db      481 CAAACTTATTGTCAAATGAGTCCCAATGTCCAGATTTTGGAACTCAACCCCTATGAACAATGG 540
QY      582 AlaIleGlnMetGluLysLysAlaAlaLysLysGlyAsnArgLysGluArgValCysAla 601
Db      541 GCCATTCAAATGGAAAAAAGCTGCAAAAGAGGAAATCGCAAGAACGTGTTTGTGCA 600
QY      602 GluHisLeuArgLysTyrAsnGluAla 610
Db      601 GAACATTTGAGGAAGTACAATGAGGCC 627
```

Search completed: January 28, 2005, 02:20:53
Job time : 6350 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 26, 2005, 19:59:28 ; Search time 216 Seconds
(without alignments)
2730.366 Million cell updates/sec

Title: US-09-515-363C-2
Perfect score: 5311
Sequence: 1 MSGYSTDENFRYLISCFRA.....LPITFPNLDYSECCLFSD 1025

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_02:
1: uniprot_sprot:
2: uniprot_trembl:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5311	100.0	1025	2 Q9H3G6	Q9h3g6 homo sapien
2	5285	99.5	1025	2 Q9BYX4	Q9byx4 homo sapien
3	4199	79.1	1025	2 Q8R5F7	Q8r5f7 mus musculu
4	4194	79.0	1025	2 Q8K5C7	Q8k5c7 mus musculu
5	3954	74.4	772	2 Q6DC96	Q6dc96 homo sapien
6	2673	50.3	693	2 Q8BZ01	Q8bz01 mus musculu
7	2633	49.6	683	2 Q8BYC9	Q8byc9 mus musculu
8	2456	46.2	468	2 Q96MX8	Q96mx8 homo sapien
9	2271.5	42.8	514	2 Q99KS4	Q99ks4 mus musculu
10	2059.5	38.8	467	2 Q8R144	Q8r144 mus musculu
11	1505	28.3	682	2 Q6GNI3	Q6gni3 xenopus lae
12	1356	25.5	678	1 LGP2_MOUSE	Q99j87 mus musculu
13	1352	25.5	678	1 LGP2_HUMAN	Q96c10 homo sapien
14	1122	21.1	244	2 Q9D2Z5	Q9d2z5 mus musculu
15	1086.5	20.5	925	2 Q95786	Q95786 homo sapien
16	1071	20.2	221	2 Q86X56	Q86x56 homo sapien
17	1035	19.5	926	2 Q6Q899	Q6q899 mus musculu
18	1035	19.5	926	2 AAS59532	Aas59532 mus muscu
19	1018	19.2	940	2 Q9GLV6	Q9glv6 sus scrofa
20	687	12.9	1037	2 Q44165	Q44165 caenorhabdi
21	678	12.8	143	2 Q8VE79	Q8ve79 mus musculu
22	574.5	10.8	1119	2 Q93413	Q93413 caenorhabdi
23	574.5	10.8	1119	2 CAB02082	Cab02082 caenorhab
24	555.5	10.5	472	2 Q8C7T2	Q8c7t2 mus musculu
25	516	9.7	398	2 Q9NT04	Q9nt04 homo sapien
26	516	9.7	620	2 Q17545	Q17545 caenorhabdi
27	453	8.5	752	2 Q9V1Z5	Q9v1z5 pyrococcus
28	441.5	8.3	650	2 O59524	O59524 pyrococcus
29	431.5	8.1	1883	2 Q8LMR2	Q8lmr2 oryza sativ
30	417	7.9	741	2 O28814	O28814 archaeoglob
31	408.5	7.7	1584	2 Q7S8J7	Q7s8j7 neurospora

32	406	7.6	1909	1 DICE_ARATH	Q9sp32 arabidopsis
33	402.5	7.6	1374	1 DCR1_SCHPO	Q09884 schizosacch
34	395.5	7.4	764	2 Q8TZH8	Q8tzh8 pyrococcus
35	392.5	7.4	778	1 YF05_METJA	Q58900 methanococc
36	391	7.4	410	2 Q8C5I3	Q8c5i3 mus musculu
37	390.5	7.4	752	2 Q6LXF6	Q6lxf6 methanococc
38	390.5	7.4	752	2 CAF30951	Caf30951 methanoco
39	389	7.3	821	2 Q8TUH1	Q8tuh1 methanosarc
40	384	7.2	1604	2 Q7XQ14	Q7xql4 oryza sativ
41	368	6.9	864	2 Q8PX35	Q8px35 methanosarc
42	365.5	6.9	837	2 Q86L44	Q86l44 dictyosteli
43	357	6.7	1912	1 DICE_HUMAN	Q9upy3 homo sapien
44	354	6.7	1906	1 DICE_MOUSE	Q8r418 mus musculu
45	354	6.7	1923	2 Q6TUI4	Q6tui4 bos taurus

ALIGNMENTS

RESULT 1

Q9H3G6

ID Q9H3G6 PRELIMINARY; PRT; 1025 AA.

AC Q9H3G6;

DT 01-MAR-2001 (TremBLrel. 16, Created)

DT 01-MAR-2001 (TremBLrel. 16, Last sequence update)

DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)

DE Melanoma differentiation associated protein-5.

GN Name=MDA5;

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Melanoma;

RX MEDLINE=21664412; PubMed=11805321;

RA Kang D., Gopalakrishnan R.V., Wu Q., Jankowsky E., Pyle A.M.,

RA Fisher P.B.;

RT "mda-5: An interferon-inducible putative RNA helicase with double-

RT stranded RNA-dependent ATPase activity and melanoma growth-suppressive

RT properties.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:637-642(2002).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Melanoma;

RA Kang D.-C., Fisher P.B.;

RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF095844; AAG34368.1; -.

DR HSSP; P10081; 1FUK.

DR GO; GO:0005524; F:ATP binding; IEA.

DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.

DR GO; GO:0003676; F:nucleic acid binding; IEA.

DR GO; GO:0015668; F:type III site-specific deoxyribonuclease ac. .; IEA.

DR GO; GO:0009307; P:DNA restriction; IEA.

DR InterPro; IPR001410; DEAD.

DR InterPro; IPR011029; DEATH-like.

DR InterPro; IPR001650; Helicase_C.

DR InterPro; IPR006935; ResIII.

DR Pfam; PF00271; Helicase_C; 1.

DR Pfam; PF04851; ResIII; 1.

DR SMART; SM00487; DEXDC; 1.

DR SMART; SM00490; HELICc; 1.

KW ATP-binding; Helicase; Hydrolase.

SQ SEQUENCE 1025 AA; 116687 MW; 48BE75491D863741 CRC64;

Query Match 100.0%; Score 5311; DB 2; Length 1025;
Best Local Similarity 100.0%; Pred. No. 3.1e-256;
Matches 1025; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSGYSTDENFRYLISCFRVRVKMYIQVEPVLDTLFLPAEYKEQIQRTVATSGNMQAVE 60
DB 1 MSGYSTDENFRYLISCFRVRVKMYIQVEPVLDTLFLPAEYKEQIQRTVATSGNMQAVE 60

```
QY 61 LLLSTLEKGVWHLGWTREFVEALRRRTGSPLAARYMNPDLTDLSPSPFENAHDEYQLQLNL 120
|
|
|
Db 61 LLLSTLEKGVWHLGWTREFVEALRRRTGSPLAARYMNPDLTDLSPSPFENAHDEYQLQLNL 120
|
|
|
QY 121 LQPTLVDKLLVRDVLDKCMEEEELLTIEDNRNRIAAENNGNESGVRELLKRIVQKENWFSA 180
|
|
|
Db 121 LQPTLVDKLLVRDVLDKCMEEEELLTIEDNRNRIAAENNGNESGVRELLKRIVQKENWFSA 180
|
|
|
QY 181 FLNVLRTGNNELVQELTGSDCSNAEIEENLSQVDGPQVEEQLLSTTVQPNLEKEVWGM 240
|
|
|
Db 181 FLNVLRTGNNELVQELTGSDCSNAEIEENLSQVDGPQVEEQLLSTTVQPNLEKEVWGM 240
|
|
|
QY 241 ENNSSSESFADSSVSESDTSLAEGSVCLDESGLHNSNMGSDSGTMGSDSDENVAARA 300
|
|
|
Db 241 ENNSSSESFADSSVSESDTSLAEGSVCLDESGLHNSNMGSDSGTMGSDSDENVAARA 300
|
|
|
QY 301 SPEPELQLRPYQMEVAQPALEGKNIIICLPTGSGKTRVAVYIAKDHLDKKKASEPGKVI 360
|
|
|
Db 301 SPEPELQLRPYQMEVAQPALEGKNIIICLPTGSGKTRVAVYIAKDHLDKKKASEPGKVI 360
|
|
|
QY 361 VLVNKVLLVEQLFRKEFQPFLLKKWYRVIGLSGDTQLKISFPEVVKSCDIIISTAQILENS 420
|
|
|
Db 361 VLVNKVLLVEQLFRKEFQPFLLKKWYRVIGLSGDTQLKISFPEVVKSCDIIISTAQILENS 420
|
|
|
QY 421 LLNLENGEDAGVQLSDFSLIIIDECHHTNKEAVYNNIMRHYLMQKLKNNRLKKNKPVIP 480
|
|
|
Db 421 LLNLENGEDAGVQLSDFSLIIIDECHHTNKEAVYNNIMRHYLMQKLKNNRLKKNKPVIP 480
|
|
|
QY 481 LPQILGLTASPGVGGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQIQEPCCKFA 540
|
|
|
Db 481 LPQILGLTASPGVGGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQIQEPCCKFA 540
|
|
|
QY 541 IADATREDPFKEKLEIMTRIQTQYCOMSPMSDFGTQPYEQWAIQMEKKAACKGNKERV 600
|
|
|
Db 541 IADATREDPFKEKLEIMTRIQTQYCOMSPMSDFGTQPYEQWAIQMEKKAACKGNKERV 600
|
|
|
QY 601 AEHLRKYNEALQINDTIRMIDAYTHLETFYNEEKDKKFAVIEDDSDEGGDDEYCDGDEDE 660
|
|
|
Db 601 AEHLRKYNEALQINDTIRMIDAYTHLETFYNEEKDKKFAVIEDDSDEGGDDEYCDGDEDE 660
|
|
|
QY 661 DDLKKPLKLDLTDRLMTLFFENNMKLKRLAENPEYENEKLTCLRNTIMEQYTRTEESAR 720
|
|
|
Db 661 DDLKKPLKLDLTDRLMTLFFENNMKLKRLAENPEYENEKLTCLRNTIMEQYTRTEESAR 720
|
|
|
QY 721 GIIFTKTROSAYALSQWITENEKFAEVGVKAHHLIGAGHSSSEFKPMTQNEQKEVISKFT 780
|
|
|
Db 721 GIIFTKTROSAYALSQWITENEKFAEVGVKAHHLIGAGHSSSEFKPMTQNEQKEVISKFT 780
|
|
|
QY 781 GKINLLIATTVAEEGLDIKECNIVIRYGLVTNEIAMVQARGARADESTYVLVAHSGSGV 840
|
|
|
Db 781 GKINLLIATTVAEEGLDIKECNIVIRYGLVTNEIAMVQARGARADESTYVLVAHSGSGV 840
|
|
|
QY 841 IEHETVNDFREKMYKAIHCVQNMKPEEYAHKILELQMOSIMEKKMKTGRNIAKHYNKP 900
|
|
|
Db 841 IEHETVNDFREKMYKAIHCVQNMKPEEYAHKILELQMOSIMEKKMKTGRNIAKHYNKP 900
|
|
|
QY 901 SLITFLCKNCSVLACSGEDIHVIEKMHVNMTPFEKELYIVRENKALQKKCADIYQINGEI 960
|
|
|
Db 901 SLITFLCKNCSVLACSGEDIHVIEKMHVNMTPFEKELYIVRENKALQKKCADIYQINGEI 960
|
|
|
QY 961 ICKCGQAWGTMVHKGLDLPCLKIRNFVVVFKNNSTKKQYKKWVELPITFPNLDYSECL 1020
|
|
|
Db 961 ICKCGQAWGTMVHKGLDLPCLKIRNFVVVFKNNSTKKQYKKWVELPITFPNLDYSECL 1020
|
|
|
1021 FSDED 1025
|
|
|
1021 FSDED 1025
|
|
|
RESULT 2
Q9BYX4
ID Q9BYX4 PRELIMINARY; PRT; 1025 AA.
AC Q9BYX4;
```

```
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE RNA helicase-DEAD box protein RH116.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=14645903;
RA Cocude C., Truong M.-J., Billaut-Mulot O., Delsart V., Darcissac E.,
RA Capron A., Mouton Y., Bahr G.M.;
RT "A novel cellular RNA helicase, RH116, differentially regulates cell
RT growth, programmed cell death and human immunodeficiency virus type 1
RT replication.";
RL J. Gen. Virol. 84:3215-3225 (2003).
DR EMBL; AY017378; AAG54076.1; -.
DR HSSP; P10081; 1FUK.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0015668; F:type III site-specific deoxyribonuclease ac. . .; IEA.
DR GO; GO:0009307; P:DNA restriction; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011029; DEATH like.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR006935; ResIII.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF04851; ResIII; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELICC; 1.
KW ATP-binding; Helicase; Hydrolase.
SQ SEQUENCE 1025 AA; 116671 MW; 6B4F3D14E88A7D6D CRC64;

Query Match 99.5%; Score 5285; DB 2; Length 1025;
Best Local Similarity 99.5%; Pred. No. 6.1e-255;
Matches 1020; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MSGYSTDENFRYLISCFRARKVMIQVEPVLDTFLPAEVKEIQIRTVATSGNQAVE 60
|
|
|
Db 1 MSGYSTDENFRYLISCFRARKVMIQVEPVLDTFLPAEVKEIQIRTVATSGNQAVE 60
|
|
|
QY 61 LLLSTLEKGVWHLGWTREFVEALRRRTGSPLAARYMNPDLTDLSPSPFENAHDEYQLQLNL 120
|
|
|
Db 61 LLLSTLEKGVWHLGWTREFVEALRRRTGSPLAARYMNPDLTDLSPSPFENAHDEYQLQLNL 120
|
|
|
QY 121 LQPTLVDKLLVRDVLDKCMEEEELLTIEDNRNRIAAENNGNESGVRELLKRIVQKENWFSA 180
|
|
|
Db 121 LQPTLVDKLLVRDVLDKCMEEEELLTIEDNRNRIAAENNGNESGVRELLKRIVQKENWFSA 180
|
|
|
QY 181 FLNVLRTGNNELVQELTGSDCSNAEIEENLSQVDGPQVEEQLLSTTVQPNLEKEVWGM 240
|
|
|
Db 181 FLNVLRTGNNELVQELTGSDCSNAEIEENLSQVDGPQVEEQLLSTTVQPNLEKEVWGM 240
|
|
|
QY 241 ENNSSSESFADSSVSESDTSLAEGSVCLDESGLHNSNMGSDSGTMGSDSDENVAARA 300
|
|
|
Db 241 ENNSSSESFADSSVSESDTSLAEGSVCLDESGLHNSNMGSDSGTMGSDSDENVAARA 300
|
|
|
QY 301 SPEPELQLRPYQMEVAQPALEGKNIIICLPTGSGKTRVAVYIAKDHLDKKKASEPGKVI 360
|
|
|
Db 301 SPEPELQLRPYQMEVAQPALEGKNIIICLPTGSGKTRVAVYIAKDHLDKKKASEPGKVI 360
|
|
|
QY 361 VLVNKVLLVEQLFRKEFQPFLLKKWYRVIGLSGDTQLKISFPEVVKSCDIIISTAQILENS 420
|
|
|
Db 361 VLVNKVLLVEQLFRKEFQPFLLKKWYRVIGLSGDTQLKISFPEVVKSCDIIISTAQILENS 420
|
|
|
QY 421 LLNLENGEDAGVQLSDFSLIIIDECHHTNKEAVYNNIMRHYLMQKLKNNRLKKNKPVIP 480
|
|
|
Db 421 LLNLENGEDAGVQLSDFSLIIIDECHHTNKEAVYNNIMRHYLMQKLKNNRLKKNKPVIP 480
|
|
|
QY 481 LPQILGLTASPGVGGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQIQEPCCKFA 540
|
|
|
```

Db 481 LPQILGLTASPGVGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQIQEPCCKFA 540

QY 541 IADATREDPFKEKLEIMTRIQTQYQMSPMDSFGTQPYEQWAIQMEKKAAGNRKERV 600

Db 541 IADATREDPFKEKLEIMTRIQTQYQMSPMDSFGTQPYEQWAIQMEKKAAGNRKERV 600

QY 601 AEHLRKYNEALQINDTIRMIDAYTHLETIFYNEEKDKKFAVIEDSDSGDDEYCDGDEDE 660

Db 601 AEHLRKYNEALQINDTIRMIDAYTHLETIFYNEEKDKKFAVIEDSDSGDDEYCDGDEDE 660

QY 661 DDLKKPLKLDTRFLMTLFFENNMKRLAENPEYENEKLTCLRNTIMEQYTRTEESAR 720

Db 661 DDLKKPLKLDTRFLMTLFFENNMKRLAENPEYENEKLTCLRNTIMEQYTRTEESAR 720

QY 721 GIIFTKTRQSAAYALSQWITENEKFAEVGVKAHHLIGAGHSSEFKPMTQNEQKEVISKERT 780

Db 721 GIIFTKTRQSAAYALSQWITENEKFAEVGVKAHHLIGAGHSSEFKPMTQNEQKEVISKERT 780

QY 781 GKINLLIATTVAEEGLDIKECNIVIRYGLVTNEIAMVQARGRARADESTYVLVAHSGSV 840

Db 781 GKINLLIATTVAEEGLDIKECNIVIRYGLVTNEIAMVQARGRARADESTYVLVAHSGSV 840

QY 841 IEHETVNDFREKMMYKAHCVQNMKPEEYAHKILELQMQSIMKKMKTKRNIKHYKNP 900

Db 841 IERETVNDFREKMMYKAHCVQNMKPEEYAHKILELQMQSIMKKMKTKRNIKHYKNP 900

QY 901 SLITFLCKNCSVLACSGEDIHVIEKMHVNMTPEFKELYIVRENKALQKCADYQINGEI 960

Db 901 SLITFLCKNCSVLACSGEDIHVIEKMHVNMTPEFKELYIVRENKALQKCADYQINGEI 960

QY 961 ICKCGQAWGTMMVHKGLDLPCLKIRNFVVVFKNSTKKQYKQWVLPITFPNLDYSECL 1020

Db 961 ICKCGQAWGTMMVHKGLDLPCLKIRNFVVVFKNSTKKQYKQWVLPITFPNLDYSECL 1020

QY 1021 FSDSD 1025

Db 1021 FSDSD 1025

RESULT 3

Q8RSF7

ID Q8RSF7 PRELIMINARY; PRT; 1025 AA.

AC Q8RSF7;

DT 01-JUN-2002 (TremBLrel. 21, Created)

DT 01-JUN-2002 (TremBLrel. 21, Last sequence update)

DT 01-MAR-2004 (TremBLrel. 26, Last annotation update)

DE HELICARD.

GN Name=9130009C22Rik;

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=22009979; PubMed=12015121;

RA Kovacsovich M., Martinon F., Micheau O., Bodmer J.L., Hofmann K.,

RA Tschoep J.;

RT "Overexpression of Helicard, a CARD-containing helicase cleaved during

RT apoptosis, accelerates DNA degradation.";

RL Curr. Biol. 12:838-843(2002).

DR EMBL; AY075132; AAL77205.1; -;

DR MGD; MGI:1918836; 9130009C22Rik.

DR GO; GO:000524; F:ATP binding; IEA.

DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.

DR GO; GO:0003676; F:nucleic acid binding; IEA.

DR GO; GO:0015668; F:type III site-specific deoxyribonuclease ac...; IEA.

DR GO; GO:0009307; P:DNA restriction; IEA.

DR InterPro; IPR001410; DEAD.

DR InterPro; IPR011029; DEATH_like.

DR InterPro; IPR001650; Helicase_C.

DR InterPro; IPR006935; ResIII.

DR Pfam; PF00271; Helicase_C; 1.

DR Pfam; PF04851; ResIII; 1.

DR SMART; SM00487; DEXDc; 1.

DR SMART; SM00490; HELICc; 1.

KW ATP-binding; Helicase; Hydrolase.

SQ SEQUENCE 1025 AA; 115970 MW; 708FCAC690C9F6D8 CRC64;

Query Match 79.1%; Score 4199; DB 2; Length 1025;

Best Local Similarity 79.7%; Pred. No. 8.2e-201;

Matches 818; Conservative 75; Mismatches 131; Indels 2; Gaps 2;

QY 1 MSGYSTDENFRYLISCFRVRVMYIQVEPVLDTLTPPAEVEQIQRTVATSGNMQAVE 60

Db 1 MSIVCSAEDSFRNLILFFRPRLKMYIQVEPVLDTLTPPAEVEQIQRTVATSGNMQAVE 60

QY 61 LLLSTLEKGVWHLGWTRFVEALRRGTGSPLAARYMNPETLDPSPSFENAHDEYLQNL 120

Db 61 LLLSTLEQGWPLGTQMFVEALEHSGNPLAARYVKTTLTDLSPSSETAHDECLHLTL 120

QY 121 LQPTLVDKLLVRDVLDKMEEELLTIEDNRNIAAENNGESGVRELLKRIQKENWESA 180

Db 121 LQPTLVDKLLINDVLTCTFEKGLLTVEDNRNISAAGNSGNESGVRELLRRIQKENWES 180

QY 181 FLNVLROTGNELVQELTGSDCSESNAEINLSQVDPQVPEEQLLSTTVQPNLEKEVGM 240

Db 181 FLDVLROTGNDALFQELTGGCPEDNTDLANSSHRDGPAAECCLPAVDESSELETAWNV 240

QY 241 ENNSSESPADSSVSESDTSLAEGSVSCLDES LGHNSNMCSGDSGTMGSDSDEENV-AAR 299

Db 241 DDILPEASCTDSSVTTESDTSLAEGSVSCLDES LGHNSNMGRDSGTMGSDSDESVIQTKR 300

QY 300 ASPEPELQRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKDHLDKKKASEPGKV 359

Db 301 VSPEPELQRPYQMEVAQPALDGKNIICLPTGSGKTRVAVYITKDHLDKKKQASEGKV 360

QY 360 IVLNVKLVLEQVLFKEFPFLKQWYVIGLSGDTQLKISPEWVKSCDIIISTAQIILEN 419

Db 361 IVLNVKVMLEAQVLFKEFNPYLKKWYRIIGLSGDTQLKISPEWVKSYDVIISTAQIILEN 420

QY 420 SLLNLENGEDAGVQLSDFSLIIIDECHHTNKEAVYNNIMRHYLMQKLKNRLKKNKPVI 479

Db 421 SLLNLESGDDGVQLSDFSLIIIDECHHTNKEAVYNNIMRHYLMQKLKNRLNNDLKKQKPAI 480

QY 480 PLPQILGLTASPGVGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQIQEPCCKF 539

Db 481 PLPQILGLTASPGVGAAKKQSEAEKHILNLCANLDAFTIKTVKENLGQLKHQIKEPCCKF 540

QY 540 AIADATREDPFKEKLEIMTRIQTQYQMSPMDSFGTQPYEQWAIQMEKKAAGNRKERV 599

Db 541 VIADDTRENPFKEKLEIMASIQYQKSPMSDFGTQHYEQWAIQMEKKAAGNRKDRV 600

QY 600 CAEHLRKYNEALQINDTIRMIDAYTHLETIFYNEEKDKKFAVIEDSDSGDDEYCDGDED 659

Db 601 CAEHLRKYNEALQINDTIRMIDAYSHLETIFYTDEKEKKFAVL-NDSDKSDDEASSCNDQL 659

QY 660 EDDLKKPLKLDTRFLMTLFFENNMKRLAENPEYENEKLTCLRNTIMEQYTRTEESA 719

Db 660 KGDVKKSLKLDDETFELMNLFFDNKKMLKLAENPKYENEKLIKLRNTILEQFTRSESS 719

QY 720 RGIIFTKTRQSAAYALSQWITENEKFAEVGVKAHHLIGAGHSSEFKPMTQNEQKEVISKERV 779

Db 720 RGIIFTKTRQSTYALSQWIMENAKFAEVGVKAHHLIGAGHSSEVKPMTQTEQKEVISKERV 779

QY 780 TGKINLLIATTVAEEGLDIKECNIVIRYGLVTNEIAMVQARGRARADESTYVLVAHSGSG 839

Db 780 TGEINLLIATTVAEEGLDIKECNIVIRYGLVTNEIAMVQARGRARADESTYVLVTSSGSG 839

QY 840 VIEHETVNDFREKMMYKAHCVQNMKPEEYAHKILELQMQSIMKKMKTKRNIKHYKNV 899

Db 840 VTEREIVNDFREKMMYKAINRVQNMKPEEYAHKILELQVQSILEKKMKVRSIAKQYNDN 899

QY 900 PSLITFLCKNCSVLACSGEDIHVIEKMHVNMTPEFKELYIVRENKALQKCADYQINGE 959

Db 900 PSLITLLCKNCSMLVCSGENIHVIEKMHVNMTPEFKGLYIVRENKALQKCFADYQTNCE 959

RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,	
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,	
RA	Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,	
RA	Fahey J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,	
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,	
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,	
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,	
RA	Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,	
RA	Jones S.J., Marra M.A.;	
RT	"Generation and initial analysis of more than 15,000 full-length human	
RT	and mouse cDNA sequences.";	
RT	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).	
RL	[2]	
RN	SEQUENCE FROM N.A.	
RP	TISSUE=Uterus;	
RC	Strausberg R.;	
RA	Submitted (JUL-2004) to the EMBL/GenBank/DBBJ databases.	
RL	EMBL; BC078180; AAH78180.1; -.	
DR	Hypothetical protein.	
KW	NON_TER 772 772	
FT	SEQUENCE 772 AA; 87604 MW; DE35A4B7D4B70EBC CRC64;	
SQ		
Query Match 74.4%; Score 3954; DB 2; Length 772;		
Best Local Similarity 99.5%; Pred. No. 9e-189;		
Matches 768; Conservative 3; Mismatches 1; Indels 0; Gaps 0;		
Qy	1 MSGYSTDENFRYLISCFRARVKMYIQVEPVDYLTFLPAEVEKEIQRTVATSGNMQAVE 60	
Db		
Qy	61 LLLSTLEKGVWHLGWTREFVEALRRTGSPLAARYMNPDLTLPSPFENAHDEYLQLNL 120	
Db		
Qy	61 LLLSTLEKGVWHLGWTREFVEALRRTGSPLAARYMNPDLTLPSPFENAHDEYLQLNL 120	
Db		
Qy	121 LQPTLVDKLLVRDVLDKMEEELLTIEDNRPIAAENNGNESGVRELLKRIVOKENWFS 180	
Db		
Qy	121 LQPTLVDKLLVRDVLDKMEEELLTIEDNRPIAAENNGNESGVRELLKRIVOKENWFS 180	
Db		
Qy	181 FLNVLRTGNLQVLTGSDCSESNAEINLSQVDGPQVEEQLLSTTVQPNLEKEVWGM 240	
Db		
Qy	181 FLNVLRTGNLQVLTGSDCSESNAEINLSQVDGPQVEEQLLSTTVQPNLEKEVWGM 240	
Db		
Qy	241 ENNSSESSFADSSVSVESDTSLAEGSVSCLDLSEGLHNSNMGSDSGTGMGSDSDEENVAARA 300	
Db		
Qy	301 SPEPELQLRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKHLDKKKKASEPGKVI 360	
Db		
Qy	301 SPEPELQLRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKHLDKKKKASEPGKVI 360	
Db		
Qy	361 VLNVKVLVLEQLFRKEFPFELKKWYRVIGLSGDTQLKISFPFVVKSCDIIISTAQILENS 420	
Db		
Qy	361 VLNVKVLVLEQLFRKEFPFELKKWYRVIGLSGDTQLKISFPFVVKSCDIIISTAQILENS 420	
Db		
Qy	421 LNLNLENGEDAGVQLSDPFLIIIDECHHTNKEAVYNNIMRHYLMQKLKNNRLKKENKPVIP 480	
Db		
Qy	481 LPQILGLTASPGVGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQIQEPCCKKFA 540	
Db		
Qy	541 IADATREDPFKEKLLEIMTRIQTQYQMSPSDFGTQPYEQWAIQMEKKAAGNKRKERV 600	
Db		
Qy	541 IADATREDPFKEKLLEIMTRIQTQYQMSPSDFGTQPYEQWAIQMEKKAAGNKRKERV 600	
Db		
Qy	601 AEHLRKYNEALQINDTIRMIDAYTHLETFTYNEEKDKKFAVIEDSDDEGDDDEYCDGDEDE 660	
Db		
Qy	661 DDLKKPLKLDDETRFLMTLFFENNKMRLKRLAENPEYENEKLTCLRNTIMEQYTRTEESAR 720	
Db		
Qy	661 DDLKKPLKLDDETRFLMTLFFENNKMRLKRLAENPEYENEKLTCLRNTIMEQYTRTEESAR 720	
Db		

Qy	721 GIIFTKTRQSAVALSQWITENEKFAEVGKKAHHLIGAGHSSEFKPMTQNEQK 772	
Db		
Qy	721 GIIFTKTRQSAVALSQWITENEKFAEVGKKAHHLIGAGHSSEFKPMTQKKKK 772	
Db		
RESULT 6		
Q8BZ01		
ID	Q8BZ01 PRELIMINARY; PRT; 693 AA.	
AC	Q8BZ01;	
DT	01-MAR-2003 (TrEMBLrel. 23, Created)	
DT	01-MAR-2003 (TrEMBLrel. 23, Last sequence update)	
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)	
DE	Mus musculus adult female vagina cDNA, RIKEN full-length enriched	
DE	library, clone:9930105B04 product:similar to MELANOMA DIFFERENTIATION	
DE	ASSOCIATED PROTEIN-5 (Fragment).	
GN	Name=9130009C22Rik;	
OS	Mus musculus (Mouse).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
OX	NCBI_TaxID=10090;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=C57BL/6J; TISSUE=Vagina;	
RX	MEDLINE=99279253; PubMed=10349636;	
RA	Carninci P., Hayashizaki Y.;	
RT	"High-efficiency full-length cDNA cloning.";	
RL	Meth. Enzymol. 303:19-44(1999).	
RN	[2]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=C57BL/6J; TISSUE=Vagina;	
RX	MEDLINE=21085660; PubMed=11217851;	
RA	RIKEN FANTOM Consortium;	
RT	"Functional annotation of a full-length mouse cDNA collection.";	
RL	Nature 409:685-690(2001).	
RN	[3]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=C57BL/6J; TISSUE=Vagina;	
RA	The FANTOM Consortium,	
RA	the RIKEN Genome Exploration Research Group Phase I & II Team;	
RT	"Analysis of the mouse transcriptome based on functional annotation of	
RT	60,770 full-length cDNAs.";	
RL	Nature 420:563-573(2002).	
RN	[4]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=C57BL/6J; TISSUE=Vagina;	
RX	MEDLINE=20499374; PubMed=11042159;	
RA	Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,	
RA	Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;	
RT	"Normalization and subtraction of cap-trapper-selected cDNAs to	
RT	prepare full-length cDNA libraries for rapid discovery of new genes.";	
RL	Genome Res. 10:1617-1630(2000).	
RN	[5]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=C57BL/6J; TISSUE=Vagina;	
RX	MEDLINE=20530913; PubMed=11076861;	
RA	Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,	
RA	Konno H., Akiyama J., Nishi K., Kitsunai T., Tashiro H., Itoh M.,	
RA	Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,	
RA	Yamanoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,	
RA	Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,	
RA	Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,	
RA	Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;	
RT	"RIKEN integrated sequence analysis (RISA) system-384-format	
RT	sequencing pipeline with 384 multicapillary sequencer.";	
RL	Genome Res. 10:1757-1771(2000).	
RN	[6]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=C57BL/6J; TISSUE=Vagina;	
RA	Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,	
RA	Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,	
RA	Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,	
RA	Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,	

RA	Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,	
RA	Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,	
RA	Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,	
RA	Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,	
RA	Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,	
RA	Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,	
RA	Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.,	
RL	Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.	
DR	EMBL; AK037057; BAC29687.1; -	
DR	MGD; MGI:1918836; 9130009C22Rik.	
DR	GO; GO:0005524; F:ATP binding; IEA.	
DR	GO; GO:0008026; F:ATP-dependent helicase activity; IEA.	
DR	GO; GO:0015668; F:type III site-specific deoxyribonuclease ac. . .; IEA.	
DR	GO; GO:0009307; P:DNA restriction; IEA.	
DR	InterPro; IPR001410; DEAD.	
DR	InterPro; IPR011029; DEATH like.	
DR	InterPro; IPR006935; ResIII.	
DR	Pfam; PF04851; ResIII; 1.	
DR	SMART; SM00487; DEXDC; 1.	
FT	NON TER 693	
SQ	SEQUENCE 693 AA; 77756 MW; BB1B458B5EE019D CRC64;	
Query Match 50.3%; Score 2673; DB 2; Length 693;		
Best Local Similarity 75.6%; Pred. No. 5.4e-125;		
Matches 525; Conservative 60; Mismatches 107; Indels 2; Gaps 2;		
QY	1 MSGYSTDENFRYLISCFRARVKMYIQVEPVLDYLTFLPAEYKEQIORTVATSGNMQAVE 60	
Db	1 MSIVCSAEDSPRNLIILFFRPRLKMYIQVEPVLDHLIFLSAETKEQILKKINTCGNTSAAE 60	
QY	61 LLLSTLEKGVHGLGWTRFVEALRRRTGSPPLAARYMNPDLTLPSPSFENAHDEYLQLNL 120	
Db	61 LLLSTLEQGWPLGWTQMFVEALEHSGNPLAARYVKPTLTDLPSSETAHDECLHLLTL 120	
QY	121 LQPTLVKLVVRDLDKCMEEELLTIEDNRNIAAENNGESGVRELLKRIYQKNWPSA 180	
Db	121 LQPTLVKLLINDVLTCTFEKGLLTVEDNRNISAAGNSGNSGVRELLRRIYQKNWFST 180	
QY	181 FLNVLRQTGNNELVQELTGSDCSESNABIEIENLSQVDGPQVEFQLLSTTVQPNLEKEVGM 240	
Db	181 FLDVLRQTGNDALEQLTGGCPEDNTDLANSRRDGPANECLLPAVDESSLETAWNV 240	
QY	241 ENNSSESSFADSSVSVSESDTSLAEGSVCLDESGLHNSNMGSDSGTMGSDSDEENV-AAR 299	
Db	241 DDILPEASCTDSSVTTESDTSLAEGSVCFDESGLHNSNMGRDSGTMGSDSDESIVQTKR 300	
QY	300 ASPEPELQLRPYQMEVAQPALEGKNIICLPTGSGKTRVAVVIAKDHLDKKKASBPQKV 359	
Db	301 VSPEPELQLRPYQMEVAQPALDGKNIICLPTGSGKTRVAVVITKDHLDKKKQASESGKV 360	
QY	360 IVLVNKKVLLVEQLFRKEFPQFLKKWYRVIGLSGDTQLKISFPEVVKSCDIIISTAQILEN 419	
Db	361 IVLVNKMVLAEQLFKEFBNPYLKKWYRIIGLSGDTQLKISFPEVVKSYDVIISTAQILEN 420	
QY	420 SLLNLENGEDAGVQLSDFSLIIIDECHHTNKEAVYNNIMRHYLMQKLKNNRLKKENKPVI 479	
Db	421 SLLNLESGDDDGVLSDFSLSIIIDECHHTNKEAVYNNIMRHYLKQKLNNNDLKKQNKPAI 480	
QY	480 PLPQILGLTASPGVGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQIQEPCCKF 539	
Db	481 PLPQILGLTASPGVGAKKQSEAEKHILNCANLDAFTIKTVKENLGQLKHQIKEPCCKF 540	
QY	540 AIADATREDPFKEKLEIMTRIQTYCQSPMSDFGTQPYEQWAIQMEKKAACKGNRKERV 599	
Db	541 VIADDTRENPFKEKLEIMASIQTYCQSPMSDFGTQHYEQWAIQMEKKAACKGNRKDRV 600	
QY	600 CAEHLRKYNALQINDTIRMIDAYTHLETTFNEEKDKKFAVIEDSDSGDDDEYCDGDED 659	
Db	601 CAEHLRKYNALQINDTIRMIDAYSHLETFTYDEKEKKFAVL-NDSDKSDDEASSCNDQL 659	
QY	660 EDDLKPKLKLDETRFLMTLFFENNKMCLKLAEN 693	
Db	660 KGDVKKSLKLDETRFLMNLFFDNKKMLKLAEN 693	

RESULT 7	
Q8BYC9	PRELIMINARY; PRT; 683 AA.
ID	Q8BYC9
AC	Q8BYC9;
DT	01-MAR-2003 (TrEMBLrel. 23, Created)
DT	01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE	Mus musculus 0 day neonate thymus cDNA, RIKEN full-length enriched
DE	library, clone:A430105A06 product:similar to MELANOMA DIFFERENTIATION
DE	ASSOCIATED PROTEIN-5 (Fragment).
GN	Name=9130009C22Rik;
OS	Mus musculus (Mouse).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX	NCBI_TaxID=10090;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=C57BL/6J; TISSUE=Thymus;
RX	MEDLINE=99279253; PubMed=10349636;
RA	Carninci P., Hayashizaki Y.;
RT	"High-efficiency full-length cDNA cloning.";
RL	Meth. Enzymol. 303:19-44(1999).
RN	[2]
RP	SEQUENCE FROM N.A.
RC	STRAIN=C57BL/6J; TISSUE=Thymus;
RX	MEDLINE=21085660; PubMed=11217851;
RA	RIKEN FANTOM Consortium;
RT	"Functional annotation of a full-length mouse cDNA collection.";
RL	Nature 409:685-690(2001).
RN	[3]
RP	SEQUENCE FROM N.A.
RC	STRAIN=C57BL/6J; TISSUE=Thymus;
RA	The FANTOM Consortium,
RA	the RIKEN Genome Exploration Research Group Phase I & II Team;
RT	"Analysis of the mouse transcriptome based on functional annotation of
RT	60,770 full-length cDNAs.";
RL	Nature 420:563-573(2002).
RN	[4]
RP	SEQUENCE FROM N.A.
RC	STRAIN=C57BL/6J; TISSUE=Thymus;
RX	MEDLINE=20499374; PubMed=11042159;
RA	Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA	Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT	"Normalization and subtraction of cap-trapper-selected cDNAs to
RT	prepare full-length cDNA libraries for rapid discovery of new genes.";
RL	Genome Res. 10:1617-1630(2000).
RN	[5]
RP	SEQUENCE FROM N.A.
RC	STRAIN=C57BL/6J; TISSUE=Thymus;
RX	MEDLINE=20530913; PubMed=11076861;
RA	Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA	Konno H., Akiyama J., Nishi K., Kitsunai T., Tashiro H., Itoh M.,
RA	Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA	Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA	Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA	Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA	Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT	"RIKEN integrated sequence analysis (RISA) system-384-format
RT	sequencing pipeline with 384 multicapillary sequencer.";
RL	Genome Res. 10:1757-1771(2000).
RN	[6]
RP	SEQUENCE FROM N.A.
RC	STRAIN=C57BL/6J; TISSUE=Thymus;
RA	Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA	Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA	Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA	Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA	Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA	Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA	Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA	Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,

RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Mix FVB/N;
RC TISSUE=Mammary tumor. WAP-TGF alpha model. 7 months old;
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC025508; AAH25508.1; -.
DR MGD; MGI:1918836; 9130009C22Rik.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR001650; Helicase C.
DR Pfam; PF00271; Helicase_C; 1.
DR SMART; SM00490; HELIC_C; 1.
SQ SEQUENCE 467 AA; 53965 MW; F3B0D976778F0442 CRC64;

Query Match 38.8%; Score 2059.5; DB 2; Length 467;
Best Local Similarity 84.6%; Pred. No. 1.2e-94;
Matches 396; Conservative 28; Mismatches 43; Indels 1; Gaps 1;

QY 558 MTRIQTQYQMSPMDFGTQPYEQWAIQMEKKAAGKGNRKERVCAEHLRKYNEALQINDTI 617
Db 1 MASITQYQKSPMSDFGTQHYEQWAIQMEKKAAGKGNRKDRVCAEHLRKYNEALQINDTI 60
QY 618 RMDIAYTHLETFFYNEEKDKFAVIEDSDSDEGGDEYCDGDEDEDLKKPLKLDETDRFLM 677
Db 61 RMDIAHSHLETFFYTDKEKKFAVL-NDSDESDDASSCNDQLKGDVKKSLKLDETDEFLM 119
QY 678 TLFFENNMKLRLAENPEYENEKLTCLRNTIMEQYTRTEESARGIIFTKTRQSAYALSQW 737
Db 120 NLFFDNKKMLKGLAENPKYENEKLIKLRNTILEQFTRSESSRGIIFTKTRQSTYALSQW 179
QY 738 ITENEKFAEVGVKAHHLICAGHSSEFKPMTQNEQKEVISKFRGTGKINLLIATTVAEEGLD 797
Db 180 IMENAKFAEVGVKAHHLICAGHSSEVKPMTQTEQKEVISKFRGTGEINLLIATTVAEEGLD 239
QY 798 IKECNIVIRYGLVTNEIAMVQARGRARESTYVLVAHSGSGVIEHETVNDFREKMYKA 857
Db 240 IKECNIVIRYGLVTNEIAMVQARGRARESTYVLVTSSGSGVTEREIVNDFREKMYKA 299
QY 858 IHCVQNMKPEEYAHKILELQMQSIMKKMTKRNIAKHYNPNPSLITFLCKNCVSLACSG 917
Db 300 INRVQNMKPEEYAHKILELQVQSILEKKMKVKRSIAKQYNDNPSLITLCKNCMLVCSG 359
QY 918 EDIHVIEKMHVNMTPPEFKELYIVRENKALQKKADYQINGEIIICKCGQAWGTMVHKGL 977
Db 360 ENIHVIEKMHVNMTPPEFKGLYIVRENKALQKKFADYQTNGEIICKCGQAWGTMVHKGL 419
QY 978 DLPCLKIRNFVVFKNNSTKKQYKQKVELPITFPNLDYSECCCLFSDSD 1025
Db 420 DLPCLKIRNFVVFNKNSPKQYKQKVELPIRFPDLDYSEYCLYSDED 467

RESULT 11
Q6GNI3
ID Q6GNI3 PRELIMINARY; PRT; 682 AA.
AC Q6GNI3;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RX MEDLINE=22341132; PubMed=12454917;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RA Klein S., Strausberg R.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC073528; AAH73528.1; -.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR006935; ResIII.
DR Pfam; PF00270; DEAD; 1.
DR Pfam; PF00271; Helicase C; 1.
DR Pfam; PF04851; ResIII; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC_C; 1.
KW ATP-binding; Helicase; Hydrolase; Hypothetical protein.
SQ SEQUENCE 682 AA; 79244 MW; 5BF182DB298BCC4C CRC64;

Query Match 28.3%; Score 1505; DB 2; Length 682;
Best Local Similarity 43.9%; Pred. No. 8.8e-67;
Matches 313; Conservative 128; Mismatches 222; Indels 50; Gaps 9;

QY 306 LQLRPYQMEVAQPALEGKNIICLPTSGKTRVAVYIAKDHLDKKKASEPGKVIVLVNK 365
Db 1 MELHDYQWEVIGPALEGKNIICLPTGAGKTRAAALYVAMRHLEMKRNA----KVCLMVNK 56
QY 366 VLLVEQLFRKEFQFLKKWYRVIGLSGDTQLKISFPVEVVKSCDIISTAQILENSLLNLE 425
Db 57 VHLVDQHFSNEFPHLKDKYKVVAISGDEHKCFFAELVQNNNDVICTAQILQNALSS-- 114
QY 426 NGEDAGVQLSDFSLIIIDECHHTNKEAVYNNIMRHYLMQKLNRLKKNKPKVIPLPQIL 485
Db 115 SSEEIHVELTDTLLIIDECHHTHKDGVYNKLMGYLERKITQKG-----KLPPQIL 165
QY 486 GLTASPGVGGATKQAKAEHILKLCANLDAFTIKTVKENLDQLKNQIQEPCKKFAIADAT 545
Db 166 GLTASPGTGRATSPFEKAEHILQICANLDTWRIMSAEVHREDLEAKAKQPNKQYDLVTER 225

QY 546 REDPFKEKLEIMTRIQTQYCOMSPM--SDFGTQPYEQWATQMEKKAACKGNRKERVCAEH 603
Db 226 PRDPFGDKLKLBMKTIHEYLRTTDFCESDFGTQLYEQKVVELEKEGAVEANRMKRTCALH 285
QY 604 LRKYNEALQINDTIRMIDAYTHLETFTFNEEKDKKFAVIEDSDSDEGGDDEYCDGDEDDDL 663
Db 286 LRKYNDLSLVHDTVRMDAYELLDYQQEK-----VIRKQN----- 322
QY 664 KKPLKLDDETDRFLMTLFFENNMKLKRLAENPEYENEKLTKLRLNTIMEQYTRTEESARGII 723
Db 323 -----DPTDAFLIQLFDGNRRARLLELAQDVRFPENPKLRLKEILRDQF-QFSSGSRGII 375
QY 724 FTKTRQSAYALSQWITENEKFAEVGVKAHHLIGAGHSSEFPMTQNEQKEVISKPTGKI 783
Db 376 FTRTRQSTHSLHNWISSKHSFQIMGVKTAPLTGAGYSNQSKHMTQNEQRETIEMPRKGQL 435
QY 784 NLLIATTVAEGLDIKECNIVIRYGLVTNEIAMVQARGRARADESTYVLVAHSGSGVIEH 843
Db 436 NLLISTVAEGLDIPQCNIVRVRYGLMTNEISMVQARGRARHEDSCYSFLAKIGGKEIRR 495
QY 844 ETVNDFREKMYKAIHCVQNMPKEEYAHKILELQMOSIMEKMKTKRNIAKHYKNPNSLI 903
Db 496 EETNETLEGLMKRAIEAVQRMPEQEYQKKIKELQEESVIARKVKQAKRDQKRNTYPEQV 555
QY 904 TFLCKNCSVLACSGEDIHVIEKMHVNMTPEFKELY-IVRENKALQKCADYQINGEIIIC 962
Db 556 RFYCRCCSQAVAHGDDFTIEGTHYVNINSDFRIIYEVCSPLDFGKKMVDWTPGKIRC 615
QY 963 KCGQAWGTMMVHKGDLPLCLKIRNFVVVFKNNSTKKQYKKWVLPITFPNLDY 1015
Db 616 LCGQDWGFEMIKYKHNFPAPISVKNFVV--ETPEIKRPRYARWKDVFPFVDELNY 666
RESULT 12
LGP2_MOUSE
ID LGP2_MOUSE STANDARD; PRT; 678 AA.
AC Q99J87; Q9D1X4;
DT 01-OCT-2004 (Rel. 45, Created)
DT 01-OCT-2004 (Rel. 45, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Probable ATP-dependent helicase LGP2 (EC 3.6.1.-) (Protein D11lgp2).
GN Name=Lgp2; Synonyms=D11lgp2e;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=21100887; PubMed=11161808; DOI=10.1006/geno.2000.6433;
RA Miyoshi K., Cui Y., Riedlinger G., Lehoczky J., Zon L., Oka T.,
RA Dewar K., Hennighausen L.;
RT "Structure of the mouse stat 3/5 locus: evolution from Drosophila to
RL zebrafish to mouse.";
RL Genomics 71:150-155(2001).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC STRAIN=C57BL/6J; Tissue=Adipose tissue;
RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA Daila E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,

RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Sempke C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RP SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.
RX PubMed=11735219; DOI=10.1006/geno.2001.6661;
RA Cui Y., Li M., Walton K.D., Sun K., Hanover J.A., Furth P.A.,
RA Hennighausen L.;
RT "The Stat3/5 locus encodes novel endoplasmic reticulum and helicase-
RT like proteins that are preferentially expressed in normal and
RT neoplastic mammary tissue.";
RL Genomics 78:129-134(2001).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q99J87-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q99J87-2; Sequence=VSP_010842, VSP_010843;
CC Note=No experimental confirmation available;
CC -!- TISSUE SPECIFICITY: Highly expressed in mammary tissues. Expressed
CC in liver and testis. Expressed at lower level in spleen, embryo,
CC mammary gland and breast tumors.
CC -!- SIMILARITY: Belongs to the helicase family.
CC -!- CAUTION: Ref.2 sequence differs from that shown due to a
CC frameshift in position 141.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; AF316999; AAK15474.1; -.
DR EMBL; AF317000; AAK15475.1; -.
DR EMBL; AK021012; BAB32276.1; ALT_FRAME.

DR EMBL; BC029209; AAH29209.1; -.
DR MGD; MGI:1931560; D11Lgp2e.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR001650; Helicase_C.
DR Pfam; PF00270; DEAD; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELICC; 1.
KW Alternative splicing; ATP-binding; Coiled coil; Helicase; Hydrolase.
FT NP_BIND 24 31 ATP (Potential).
FT SITE 131 134 DECH box.
FT DOMAIN 489 546 Coiled coil (Potential).
FT VARSPLIC 188 189 LC -> VS (in isoform 2).
FT VARSPLIC 188 189 /FTID=VSP 010842.
FT VARSPLIC 190 190 Missing (in isoform 2).
FT VARSPLIC 190 190 /FTID=VSP 010843.
SQ SEQUENCE 678 AA; 76726 MW; DC42B75A3AD376A8 CRC64;

Query Match 25.5%; Score 1356; DB 1; Length 678;
Best Local Similarity 41.7%; Pred. No. 2.3e-59;
Matches 299; Conservative 129; Mismatches 237; Indels 52; Gaps 12;

QY 306 LQLRPYQMEVAQPALEGKNIICLPTGSGKTRVAVYIAKHLDKKKKASEPGKVIVLVNK 365
Db 1 MELRPYQWEVILPALEGKNIICLPTGAGKTRAAAFVAKRHLE---TVDRGKVVLVNR 56

QY 366 VLLVEQLFRKEFPFLKMYRVIGLSGDTQLKISFPEVVKSCDIIISTAQILENSLNLLE 425
Db 57 VHLVSQ-HAEFRRMLDKHTVTTLSDGMSRAGFGLMARSHDLLICTAELLQLALNSSE 115

QY 426 NGEDAGVQLSDFSLIIDECHHTNKEAVVNNIMRHYLMQKLNRLKKNKPKVIPLPQIL 485
Db 116 --EDEHVELREFSLIVDECHHTKDTVTNTILSRYLEQKLK----KAE-----PLPQVL 164

QY 486 GLTASPGVGATKQAKABEHILKLCANLDAFTIKTVKENLDQLKNQIQEPCPKKFAIADAT 545
Db 165 GLTASPGTGGATKLGQAIIDHILQLCANLDTCHIMSPKNCYSQLLMHNPCKQYDLCQRR 224

QY 546 REDPFKEKLEIMTRIQTQYCMSPM-SDFGTQPYEQWAIQMEKKAAGKGNRKERVCAEHL 604
Db 225 AQDPFGDLIKLMNQIHQLEMPDLKQKQFGTQMYEQVVQLCKDAEAGLQEQRVYALHL 284

QY 605 RKYNEALQINDTIRMIDAYTHLETFFYNEEKDKKFAVIEDDSDEGGDEYCDGDEDEDLKL 664
Db 285 RRYNDALFIHDTVRARDALDMLQDFYDRERTTKQMVRAES----- 325

QY 665 KPLKLDDETRFLMTLFFENNMKLRKLAENPEYENKLTCLRNTIMEQYTRTEESARGIIF 724
Db 326 -----WLLKLFDDHKVNLGQLAARGP-ENPKLEMLERILLKQF-GSPGHTRGIIIF 373

QY 725 TKTRQSAYALSQWITENEKFAEVGVKAAHHLIGAGHSSEFKPMTONEQKEVISKFRTGKIN 784
Db 374 TRTRQTASSLLWLRQOPCLQTVGIKQPMLIGAGNTSQSTHMTQKQQEVIQEFRDGILS 433

QY 785 LLIATTVAEEGLDIKECNIVIRYGLVTNETIAMVQARGRAREDESTYVLVAHSGSGVIEHE 844
Db 434 LLVATSVAAEEGLDIAQCNVWVRYGLLTNETISMVQARGRARGQSVYSFLATEGSRMKRE 493

QY 845 TVNDFREKMYKAIHCYQNMKPEEYAHKILELQMQSIMKKMKTKRNIKHYKKNPSLIT 904
Db 494 LTNEALEVLMKAVAQVKMDPDEFKAKIRDLQOASLVKRAARAAREIQQGFLPEHVQ 553

QY 905 FLCKNCVSLACSGEDIHVEIKMHVNMTPFKELYIVRENK-ALQKKCADYQINGEIIIC- 962
Db 554 LLCINCMVAVGYSGLRKVEGTHVNVNPNFVYTTTSQNPVVINKVFKDWRPGGTIRCS 613

QY 963 KCGQAWGTMVHKGLDLPCLKIRNFVVFKNNSTKKQYKKWVELPITFPNLDYSECC 1019
Db 614 NCGEVWGFQMIYKSVTLPLVKIGS--MLLETFRGKIQAQKWSRVFPIVFDILQDC 668

RESULT 13
LGP2_HUMAN

ID AC Q96C10; Q9HAM6; STANDARD; PRT; 678 AA.
DT 01-OCT-2004 (Rel. 45, Created)
DT 01-OCT-2004 (Rel. 45, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Probable ATP-dependent helicase LGP2 (EC 3.6.1.-) (Protein D11Lgp2 homolog).
GN Name=LGP2; Synonyms=D11LGP2E;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX PubMed=14702039; DOI=10.1038/ng1285;
RA Ota T., Suzuki Y., Nishikawa T., Ohtsuki T., Sugiyama T., Irie R., Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H., Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S., Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahara K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K., Tanai H., Kimata S., Watanabe M., Hiraoka S., Chiba Y., Ishida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T., Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O., Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N., Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S., Yoshikawa Y., Matsunawa H., Ichihara T., Shiohara N., Sano S., Moriya S., Momiya H., Satoh N., Takami S., Terashima Y., Suzuki O., Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H., Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B., Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y., Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T., Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y., Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S., Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T., Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M., Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T., Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K., Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R., Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
RT "Complete sequencing and characterization of 21,243 full-length human cDNAs.";
RL Nat. Genet. 36:40-45(2004).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D., Klausner R.D., Collins F.S., Wagner K.H., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E., RA Schmerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the helicase family.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AK021416; BAB13818.1; -.
EMBL; BC014949; AAI14949.1; -.
HSSP; Q9WY48; 1GMS.
MIM; 608588; -.
InterPro; IPR001410; DEAD.
InterPro; IPR001650; Helicase_C.
Pfam; PF00270; DEAD; 1.
Pfam; PF00271; Helicase_C; 1.
SMART; SM00487; DEXDC; 1.
SMART; SM00490; HELICC; 1.
KW ATP-binding; Coiled coil; Helicase; Hydrolase; Polymorphism.
NP BIND 24 31 ATP (Potential).
SITE 131 134 DECH box.
DOMAIN 489 546 Coiled coil (Potential).
VARIANT 425 425 Q -> R (in dbSNP:2074158).
/FTID=VAR 019394.
VARIANT 523 523 R -> Q (in dbSNP:2074160).
/FTID=VAR 019395.
CONFLICT 473 473 R -> W (in Ref. 1).
SEQUENCE 678 AA; 76612 MW; 859E1749C7313D06 CRC64;

Query Match 25.5%; Score 1352; DB 1; Length 678;
Best Local Similarity 40.7%; Pred. No. 3.7e-59;
Matches 292; Conservative 136; Mismatches 237; Indels 52; Gaps 12;

QY 306 LQLRPYQMEVAQPALEGNIIICLPTGSGKTRVAVYIAKDLDKKKKASEPGKVIIVLNK 365
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 365
1 MELRSYQWEVIMPALEGNIIICLPTGAGKTRAAAYVAKRHLE-----TVDGAKVVVLVNR 56
QY 366 VLLVEQLFRKEFQFLKKYRVVIGLSGDTQLKISFPEVVKSCDIIISTAQILENSLLNLE 425
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 425
57 VHLVTQ-HGEFRRMLDGRWTVTTLSGDMGPRAGFGLHARCHDLICTAELLQMALTSPE 115
QY 426 NGEDAGVQLSDFSLIIDECHHTNKEAVYNNIMRHLYMQKLKNNRLKKNKPVIPLPQIL 485
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 485
116 --EEHVELTVFSLIVVDECHHTKDTVYNNVMSQYLELKLQRAQ-----PLPQVL 164
QY 486 GLTASPGVGGATKQAKAEHILKLCANLDAFTIKTVKENLDQKNQIQEPCPKFAIADAT 545
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 545
165 GLTASPGTGGASKLDGAINHVLQCANLDTWCIMSPQNCPCPQLEHSQQPCQYNLCRR 224
QY 546 REDPFKEKLEIMTRIQTQYQMSQSPMS-DFGTQPYEQWAIQMEKKAAGKGNRKERVCAEHL 604
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 604
225 SQDPFGDLLKLMQIHDLHLEMPELSRRKFGTQMYEQVQVVKLSEAAALAGLQEQRVYALHL 284
QY 605 RKYNEALQINDTIRMIDAYTHLETIFYNEEKDKKFAVIEDDSDEGGDDCYDGEDDEDLK 664
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 664
285 RRYNDALLIHDTVRAVDALAAALQDFYHREHVTQTQIL-----C----- 322
QY 665 KPLKLDTRFLMTLFFENNKMFKRLAENPEYENKLTKLRTNIMEQYTRTEESARGIIF 724
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 724
323 -----AERLLALFDDRKNELAHATHGP-ENPKLEMLEKILQRFPS-SSNSPRGIIF 373
QY 725 TKTRQSAVALSQWITENEKFAEVGVKAHHLIGAGHSSEFFKPMQTONEQKEVISKFTGKIN 784
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 784
374 TRTRQSAHSLLLWLQOQGLQTVDIRAQLLIGAGNSSQSQTHMTQDQOQVEIQKFQDGTLN 433
QY 785 LLIAATTVAEGLDIKECNIVIRYGLVTNEIAMVQARGARADESTYVLVAHSGSGVIEHE 844
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 844
434 LLVATSAVEEGLDIPHNCVVVRYGLTNEISMVQARGARADQSVAFVATEGSRELKRE 493
QY 845 TVNDFREKMYKAIHCVQNMPPEYAKHILELQMQSIMEKKWTKRNIKHYKNNPSLIT 904
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 904
494 LINEALETLMQAVAAVQKMDQAEYQAKIRDLOQAALTKRAAQAQRENQROQFPPEHVQ 553

QY 905 FLCKNCSVLACSGEDIHVIEKMHVNMTPFEKELY-IVRENKALQKKCADYQINGEIIICK 963
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 963
554 LLCINCVMVAVGHGSDLRKKEGTHVNVNPNFNSYNYVSRDPVINKVFXDKWPGGVISCR 613
QY 964 -CGQAWGTMMVHKGLDLPLCKIRNFVVVFKNNSTKQYKKWVELPITFPNLDYSECC 1019
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 1019
614 NCGEVWGLQMIYKSVKLPVLKVR--MLLETPQGRQAQKWSRVFSPDFDLQHC 668

RESULT 14

Q9D2Z5 PRELIMINARY; PRT; 244 AA.
ID Q9D2Z5
AC Q9D2Z5;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Mus musculus adult male cecum cDNA, RIKEN full-length enriched library, clone:9130009C22 product:similar to MELANOMA DIFFERENTIATION ASSOCIATED PROTEIN-5 (Fragment).
GN Name=9130009C22Rik;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cecum;
RX MEDLINE=99279253; PubMed=10349636;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning."
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cecum;
RX MEDLINE=21085660; PubMed=11217851;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cecum;
RA The FANTOM Consortium;
RT the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cecum;
RX MEDLINE=20499374; PubMed=11042159;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M., Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes."
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cecum;
RX MEDLINE=20530913; PubMed=11076861;
RA Shibata K., Itoh M., Aizawa K., Nishi K., Kitsunai T., Tashiro H., Itoh M., Konno H., Akiyama J., Nishi K., Hazama M., Nishine T., Harada A., Sumi N., Ishii Y., Nakamura S., Ikegami T., Kashiwagi K., Yamamoto R., Matsumoto H., Sakaguchi S., Izawa M., Ohara E., Watahiki M., Fujiwaki S., Inoue K., Togawa Y., Tanaka T., Matsuura S., Kawai J., Okazaki Y., Ishikawa T., Ozawa K., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer."
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cecum;

QY

750

KAHLICAGHSSEFKPMTQNEQEVISKFR-TGKINLLIATTVAEEGLDIKECNIVIRYG

808

Db

661

-----LTGRGKTNTGMTLPAQKCILDFAFKASGDHNLIAATSVADEGIDIAQC�NLVILYE

716

QY

809

LVTNEIAMVQARGRARADESTYVLVAHSGSGVIEHETVNDFREKMMYKAHCVQNMKPEE

868

Db

717

YVGNVTKMIQTRGRGRARGSKCFLLT-SNAGVIEQINMYKEKMMNDISLRLQTWDEAV

775

QY

869

YAHKILELQMQSIMEKMMKTKRNIAKHY--KNNPSLITFLCKNCVSLACSGEDIHVIEKM

926

Db

776

FREKILHIQTH---EKFIRDSQEKPKPVPDKENKKL---LCRKCKALACYTADVVRVIEEC

829

QY

927

HHVNMTPEPKELYIVRENKALQKCKADYQINGEIIIC---KCGQAWGTMVHKGLDLPCLK

983

Db

830

HYTVLGDAFKECFVSRPHPK-PKQFSSFEKRAKIFCARQNCSDWDGIIHVKYKTFEIPVIK

888

QY

984

IRNFVVVFKNNSTKKQYKKW

1003

Db

889

IESFVVVEDIATGVQTLYSKW

908

Search completed: January 26, 2005, 20:14:04 /
Job time : 225 secs